Semiparametric Estimation of Willingness to Pay Distributions[1]

Mark Yuying An

Department of Economics, Duke University, Durham, NC 27708-0097.
Telephone: (919) 660-1809. Fax: (919) 684-8974. Email: man@econ.duke.edu

July, 1996.

[1] I thank Roberto Ayala, Paul Ruud and Kerry Smith for helpful discussions and Duke University for partial financial support. An electronic version of the paper can be obtained in postscript format using anonymous ftp as "pub/man/papers/spwtp.ps" on lewis.econ.duke.edu.

Abstract

The most popular survey method used in contingent valuations asks "open-ended" dichotomous choice questions. This method generates grouped or interval-censored data on respondents' willingness to pay. This paper specifies the willingness to pay distribution using the proportional hazard specification in duration analysis. This semiparametric distribution, on the one hand, controls for the effects of observed personal characteristics, and on the other, allows the shape of the distribution to be unspecified. To estimate the willingness to pay distribution from grouped data, we propose both a maximum likelihood estimation method and a minimum Chi-square method. The latter procedure applies to "many observations per cell" cases where the observable covariates are either categorical or amendable to sensible grouping. Specification tests for the proportionality assumption are proposed. The statistical inference procedures are illustrated using the data set from the San Joaquin Valley contingent valuation survey.

Key Words: Contingent Valuation, Willingness to Pay, Semiparametric Estimation, Minimum Chi-Square.

JEL Classification: C14 C25 Q20

1 Introduction

One important task both in environmental economics and in public economics is to assess the value of some non-market good. The central concept has been consumers' willingness to pay. To estimate the willingness to pay distribution, researchers often use data from contingent valuation surveys.[2] While heated debate over the validity of the contingent valuation surveys continues, it is generally agreed that — if a contingent valuation survey is conducted — "close-ended," dichotomous choice type of questions are preferred to "open-ended" dollar amount questions. From the perspective of the analyst, dichotomous choice questions generate grouped or interval-censored measurements of respondents' willingness to pay.

To estimate the willingness to pay distribution with such grouped data, researchers so far have used either parametric or nonparametric approaches.[3] The parametric approach specifies a distribution function for the willingness to pay, possibly conditional on a vector of covariates, with only a finite number of unknown parameters. The statistical inference follows a standard ordered discrete choice framework for which the maximum likelihood procedure is now routine and included in many computing packages. The nonparametric approach can be applied either using Ayer et al's (1955) *pooling adjacent violators* algorithm or Turnbull's (1974) *self-consistent* algorithm. Recently, An and Ayala (1996) recognize that there is a data grouping mechanism found in some contingent valuation surveys to which Turnbull's method does not apply. They refer to these cases as distinct bids and mixed bids, where *across-interval-censored* observations are encountered. An and Ayala present a *generalized self-consistent algorithm* that estimates the distribution function for *across-interval-censored* data.

[2] For a comprehensive review of the contingent valuation method see Cummings *et al* (1986) and Mitchell and Carson (1989). For a critical view of the method see Hausman (1993) and McFadden (1994). For an assessment of the NOAA Panel Guidelines see Carson *et al* (1996).

[3] Huang *et al* (1995) use Monte Carlo simulations to evaluate and compare the properties of the parametric and nonparametric estimators of the mean willingness to pay with binary choice data.

Neither the parametric nor the nonparametric extremes are considered fully satisfactory. The former is susceptible to severe specification error. The latter does not allow the analyst to accommodate observed covariates, an important ingredient in economic analysis. To cope with these problems, a semiparametric framework seems to be most appropriate. At first glance, this lack of activity seems odd given the fact that during the last two decades, many semiparametric estimation procedures for discrete choices have been developed in the econometrics literature.[4] These methods have not been embraced by contingent valuation researchers for at least three reasons. First, these methods are often designed for the estimation of the parameters representing the effect of covariates. The underlying distribution of the dependent variables is treated only as a nuisance part of the model. Second, with the exception of Han (1987), all these methods discuss simple threshold binary choices. Ordered multiple discrete choice models, which are the most relevant to contingent valuation studies, have not been substantial. Third, for general discrete choice models the exact distribution of the underlying distribution is identifiable only up to a scale factor. Contingent valuation surveys provide key information to identify the scale.

This paper considers a semiparametric estimation of the willingness to pay distribution with grouped data from dichotomous choice surveys. One of the contributions of this paper is that, by exploiting the similarity between the willingness to pay distribution and the statistical models for survival times, we propose to represent the willingness to pay distribution using the semiparametric proportional hazard specification. This specification, on the one hand, controls for the effects of observed heterogeneity, and on the other, allows the shape of the distribution to remain unspecified.

For statistical inference of the semiparametric willingness to pay distribution from

[4]These include the maximum score estimator of Manski (1975) and Horowitz (1992), the distribution free maximum likelihood estimator of Cosslett (1983), the maximum rank correlation estimator of Han (1987), the semiparametric least square estimator of Ichimura (1993), the sieve estimator of Klein and Spady (1993), and the iterative least squares estimator of Wang and Zhou (1995).

grouped data, we discuss both a maximum likelihood estimation procedure and a minimum Chi-square procedure. Both procedures identify and consistently estimate the same set of parameters.

The minimum Chi-square estimation procedure consists of two easy steps. In the first, one adopts some nonparametric method to consistently estimate the cell-specific survivor functions. In the second, one constructs some nonlinear transformations of the survivor function which are linear in the parameters. The model can then be estimated by performing generalized least squares to the system of seemingly uncorrelated equations.

A similar minimum Chi-square procedure has been applied to other areas such as the binary choice model (Berkson, 1953, Zellner and Lee, 1965) and duration analysis (Ryu, 1994, An 1996). In those two areas, data groupings are simple and the estimation of the cell-specific survivor function required in the first step is straightforward. In some contingent valuation surveys, the data grouping mechanism complicates the first step. The second contribution of the paper is to integrate into the minimum Chi-square procedure the available nonparametric methods of estimating the cell-specific functions with grouped data.

Just as in any other application of the minimum Chi-square procedure, the main restriction of the procedure is that it applies only to a "many observations per cell" situation in which all explanatory variables are either categorical or amendable to sensible grouping. This is a strong restriction in principle but not as severe in practice, since in many cases categorical or grouped explanatory variables are the only information in the data.

The rest of the paper is organized as follows. Section 2 begins with a brief definition of willingness to pay. This is followed by the introduction of the proportional hazard specification for the willingness to pay distribution. The date grouping mechanism encountered in contingent valuation surveys is then discussed. In section 3 we discuss the maximum likelihood estimation procedure. Identification issues are addressed with the help of the sample likelihood function. Section 4 describes the minimum Chi-square procedure. The minimum Chi-square inference proposed in this paper depends crucially on ways to consistently estimate the cell-specific survivor function. This is done in this paper by applying An and Ayala's (1996) generalized self-consistent algorithm. Section 5 investigates three important issues: specification tests for the proportional hazard assumption, estimation of mean willingness to pay and estimation of marginal effects. In section 6 both the maximum likelihood estimation procedure and the minimum Chi-square procedure are illustrated using data from a contingent valuation study conducted by Hanemann *et al* (1991). Section 7 concludes with a summary of the main points of the paper.

2 Willingness to Pay

2.1 Definition of willingness to Pay

A respondent's willingness to pay is defined as the dollar amount, Y, which equalizes the two indirect utilities with and without the provision of the non-market good in question,

$$V_1(I - Y|Z, \epsilon) = V_0(I|Z, \epsilon),$$

where I is the disposable income of the respondent; Z a vector of observed social demographic characteristics of the respondent; and ϵ a scalar variable representing uncontrolled personal characteristics. $V_1 : \mathbf{R}_+ \mapsto \mathbf{R}$ and $V_0 : \mathbf{R}_+ \mapsto \mathbf{R}$ are, respectively, the respondent's indirect utility with or without the provision of the public good. Assume that for any fixed (Z, ϵ), $V_1(x|Z, \epsilon)$ is monotonic increasing in the empirically relevant range. Then there exists an inverse function $U(x; Z, \epsilon) : \mathbf{R} \mapsto \mathbf{R}_+$ such that,

$$U(V_1(x|Z, \epsilon); Z, \epsilon) = x, \qquad x \geq 0.$$

Using this notation, the willingness to pay can be expressed as,

$$Y = I - U(V_0(I|Z, \epsilon); Z, \epsilon) = \phi(I, Z, \epsilon). \tag{1}$$

The above definition of the willingness to pay has four important implications.

(a) Since ϵ is unobserved, $Y = \phi(I, Z.\epsilon)$ is a random variable whose distribution, conditional on the observable (I,Z), is determined by the distribution of the unobservable ϵ.[5]

(b) Related to (a), a tightly specified economic theory might result in functional forms for V_1 and V_0, and hence also a functional form for ϕ. But the distribution of ϵ has to come from outside economic theory. Therefore no theory can generate the complete parametric form for the distribution of Y. This is one of the reasons why semiparametric willingness to pay distribution seems most appropriate.

(c) By definition, $V_1(x|Z, \epsilon) \geq V_0(x|Z, \epsilon)$ for any fixed (Z, ϵ) and for all $x \geq 0$. It then follows that support of the distribution of Y is the interval $[0, I] \subset \mathbf{R}_+$.

(d) Related to (c), while it is important to control for the effect of the observable (I, Z), I in general should be treated differently from Z since I also defines the support of the distribution of Y.

2.2 A Semiparametric Model

The above discussion demonstrates that the willingness to pay distribution shares many features with the duration variables in survival analysis. The distribution of a duration variable is often specified in terms of its hazard function in lieu of its density function. While the term "hazard" has a strong connotation of duration or survival, from the modeling point of view there is nothing at all special about the hazard function approach. The hazard function and the density function are one-to-one. In this paper we adopt the following proportional hazard specification for

[5] Here we assume random sampling in the sense that (I, Z) is weakly exogenously observed. This assumption is usually satisfied in practice. Otherwise, one is obligated to model the joint distribution of (Y, I, Z) instead of the conditional distribution Y given (I, Z) to construct the valid sample likelihood function for the model.

the willingness to pay distribution. Under this specification, the hazard function conditional on $X = (I, Z)$ is

$$h(y|X) \equiv -\frac{d}{dy}\log(S(y|X) = h_0(y)e^{\beta' X}1_{y\in[0,I]}. \tag{2}$$

To allow the semi-parametric structure of the model, the baseline hazard $h_0(y)$ will be left unspecified. In particular, this specification allows for asymmetry and multi-modality, probably two of the most important stylized facts found by empirical researchers.[6] The corresponding survivor function is

$$S(y|X) = \exp\left\{-e^{\beta' X}H(y)\right\}, \qquad y \geq 0, \tag{3}$$

where $H(y) = \int_0^y h_0(s)ds$ is called the integrated baseline hazard. Following point (c) the willingness to pay distribution is non-defective in the sense that $S(I|X) = 0$ for all X. A necessary and sufficient condition for this is $\lim_{y\to I} H(y) = \infty$. We will therefore maintain this implicit assumption in the sequel.

2.3 The Data Grouping Mechanism

One of the most frequently used elicitation methods in contingent valuation studies is the so called double-bounded dichotomous choice question. In this setting, each respondent answers a sequence of two questions as to whether he or she would be willing to pay some specified amounts of money to obtain the non-market good. The dollar amount used in the first question is called the original bid, denoted by B^o. The dollar amount used in the second question depends on the response given to the first question. It is a higher amount, B^h, if the answer to the first question is affirmative; and a lower amount, B^l, if the answer is negative. The answers to these two questions reveal that the respondent's willingness to pay falls into one of the four intervals $[0, B^l)$, $[B^l, B^o)$, $[B^o, B^h)$, and $[B^h, I]$, which constitute a partition of the

[6] An and Ayala (1995) demonstrate that a parametric mixture model not only fits the data better but also leads to more sensible estimates for the welfare measure, relative to a unimodal, single regime distribution.

distribution support $[0, I]$. Therefore from a random sample of N respondents, the data information available to the analyst is for each respondent n,

$$\{X_n, B_n^l, B_n^o, B_n^h, d_n\} \tag{4}$$

where $X_n = (I_n, Z_n)$ is a vector of covariates (such as income, education level, race, gender, religion, location group, etc.); $d_n \in \{1, 2, 3, 4\}$ is the respondent's answer representing "no-no," "no-yes," "yes-no," and "yes-yes" respectively.

In principle there are many ways to generate the bids (B_n^l, B_n^o, B_n^h) to be used in surveys. A common practice is to pre-specify a small number (K) of versions of the questionnaires. Each version differs by the dollar amount of each bid and is randomly assigned to the respondents. Let $b_1 < b_2 < \ldots < b_T$ be the ordered permutation of all the distinct bid amounts from the questionnaire. Define $b_0 = 0$, $b_{T+1} = \infty$.[7] Then, an answer to the double-bounded dichotomous choice question will reveal that the respondent's willingness to pay falls into an interval $Y_n \in [b_j, b_k)$ for some $j < k$. In the language of An and Ayala (1996), interval $[b_0, b_j)$ corresponds to left censoring; $[b_j, b_{T+1})$, to right censoring; $[b_j, b_{j+1})$, to interval censoring; and $[b_j, b_k)$ for $k > j + 1$, to across-interval censoring. All those cases are possible from double-bounded dichotomous choice data.

3 Maximum Likelihood Estimation

3.1 Sample Likelihood

To facilitate discussion, more notation is needed. For each individual n, define $(T + 1)(T + 2)/2$ binary indicators of the form,

$$\gamma_{nij} = 1_{y_n \in [b_i, b_j)}, \qquad 0 \leq i < j \leq T + 1, \tag{5}$$

[7] In fact we should write $b_{T+1} = I$ following the earlier discussion. In practice, $b_T << I_n$, for all n. Later on we will show that the tail of willingness to pay distribution beyond b_T is unidentified. Therefore the two treatments make no difference in terms of the statistical inference of the identifiable part of the model.

with $\gamma_{n,0,T+1}$ fixed at 0. For each n, one and only one γ_{nij} is one and the rest are zero. Let $\gamma_{ij} = \sum_{n=1}^{N} \gamma_{nij}$, the grouped data of individuals' willingness to pay can then be read from the following matrix of sufficient statistics,

$$\begin{bmatrix} \gamma_{01} & \gamma_{02} & \gamma_{03} & \cdots & \gamma_{0T} & 0 \\ - & \gamma_{12} & \gamma_{13} & \cdots & \gamma_{1T} & \gamma_{1,T+1} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ - & - & - & \cdots & \gamma_{T-1,T} & \gamma_{1,T+1} \\ - & - & - & \cdots & - & \gamma_{T,T+1} \end{bmatrix}. \tag{6}$$

In terms of the γ's, the sample log likelihood is

$$\sum_{n=1}^{N} \sum_{1=0}^{T} \sum_{j=I+1}^{T+1} \gamma_{nij} \log \left(S(b_i|X_n) - S(b_j|X_n) \right), \tag{7}$$

where $S(y|X)$ is defined in (3).

3.2 Identification

We will now demonstrate that data grouping essentially puts an upper bound on the flexibility of the continuous distribution. The log likelihood function (7) depends only on T points of the survivor function, $S(b_j|X_n)$ for $1 \leq j \leq T$. It is independent of any intermediate value. By the proportional hazard assumption, the maximum flexible functional form would be to define T extra parameters $\delta_j = H(b_j)$ for $1 \leq j \leq T$, so that the sample log likelihood function can be written as

$$l(\beta, \delta_1, \cdots, \delta_T) = \sum_{n=1}^{N} \sum_{I=0}^{T} \sum_{j=I+1}^{T+1} \gamma_{nij} \log \left(\exp\{-e^{\beta' X_n} \delta_i\} - \exp\{-e^{\beta' X_n} \delta_j\} \right). \tag{8}$$

In other words, the maximum flexible semiparametric structure which can be identified with grouped data takes a *parametric form.*

For proportional hazard duration models, Cox's (1972) maximum partial likelihood estimator provides consistent and asymptotically normal estimates for β without specification of the nuisance baseline hazard $h_0(y)$. However, Cox's procedure requires formally continuous data. Available tie-breaking approximations (e.g. Cox and Oakes, 1984) achieve reasonable precision only if ties are rare. With grouped data there seems no way to avoid the estimation of the baseline hazard. One way to achieve consistency in estimating β and to accommodate maximum flexibility of the baseline hazard is to specify it in a *piece-wise constant* form,[8]

$$h_0(y) = \sum_{j=1}^{T} c_j 1_{(b_{j-1} \leq y < b_j)}. \tag{9}$$

In this setting, $(c_1, c_2, .., c_T)$ together with β are jointly estimated with maximum likelihood estimation (Kiefer, 1988). There, too, the data grouping causes the model to be intrinsically finite-dimensional. It is easy to see that the T number of c's in (9) and the T number of δ's in (8) are exactly one-to-one. In fact,

$$\delta_j = \sum_{I=1}^{j} c_j (b_j - b_{j-1}), \qquad 1 \leq j \leq T.$$

Notice also that while the assumption (9) is a special assumption on the baseline hazard (2) that defines δ_j's, assessment of the mean value of the willingness to pay and other welfare measures calls for some interpolation of the integrated hazard function. Assumption (9) might as well be the natural choice (See Figure 1 and more detailed discussion in Section 5.2).

Maximum likelihood estimator $\hat{\theta}$ of $\theta = (\beta', \delta_1, \cdot, \delta_T)'$ is defined as the one that maximizes (8) subject to $0 \leq \delta_1 \leq \delta_j \leq \delta_T$. For this hill climbing algorithms are readily available. Statistical inference can be conducted using the first-order asymptotic normality result which states that the $\hat{\theta}$ is asymptotically normal with mean θ and a variance-covariance matrix Σ which can be consistently estimated by

$$\hat{\Sigma} = \left[\frac{\partial l_2(\hat{\theta})}{\partial \theta} \frac{\partial l_2(\hat{\theta})}{\partial \theta'} \right]^{-1}$$

at low cost.

[8] See Lancaster, 1990.

4 Minimum Chi-Square Procedure

In this section we propose an alternative estimation procedure applied to "many-observations-per-cell" data in which the observed covariates are either categorical or amendable to sensible grouping. The maintained hypothesis in this section is that there is a set of pre-specified vector values $W = \{w^1, w^2, ..., w^M\}$ such that $X_n \in W$ for all n. The whole sample is therefore divided into M cells. The observations within each cell share the same X vector. Let N^m be the size of cell w^m. That is,

$$N^m = \sum_n 1_{(X_n = w^m)}. \tag{10}$$

For all $0 \le j < k \le T+1$ and $1 \le m \le M$, let

$$\gamma_{jk}^m = \sum_{n=1}^N 1_{b_j < Y_n \le b_k} 1_{X_n = w^m}, \tag{11}$$

be the number of observations whose willingness to pay Y_n falls in the interval $(b_j, b_k]$ and whose covariates are w^m. Then, it is clear that

$$\{b_j, \gamma_{jk}^m, w^m\} \qquad m = 1, 2, ..., M, \qquad 0 \le j < k \le T+1 \tag{12}$$

constitute the sample sufficient statistics.

Exploiting the fact that the willingness to pay Y is grouped and the covariates X are categorical, define

$$S_j^m = S(b_j | w^m) = \exp\{-e^{\beta' w^m} \delta_j\}, \tag{13}$$

to ease exposition.

4.1 The Estimator

The key step in developing the minimum Chi-square estimator is to find a transformation of the cell probabilities which is linear in the parameters of interest. In the current context, consider a typical respondent in cell m. The conditional probability that $Y > b_j$ given that $Y > b_{j-1}$ is

$$S_j^m / S_{j-1}^m = \exp\left\{-e^{\beta' w^m}(\delta_j - \delta_{j-1})\right\}.$$

That is,

$$\log[-\log(S_j^m / S_{j-1}^m)] = \beta' w^m + \lambda_j \qquad j = 1, 2, ...T \tag{14}$$

where $\lambda_j = \log(\delta_j - \delta_{j-1})$.

Given $\sqrt{N}$-consistent estimators $\hat{S}_j^m$ of S_j^m,[9] define $\hat{\alpha}_j^m = \hat{S}_j^m / \hat{S}_{j-1}^m$ as the estimator for $\alpha_j^m = S_j^m / S_{j-1}^m$. For all $j = 1, 2, ..., T$ and all $m = 1, 2, .., M$, rewrite (14) as

$$\log[-\log(\hat{\alpha}_j^m)] = \beta' w^m + \lambda_j + v_j^m, \qquad j = 1, 2, ..., T \tag{15}$$

where $v_j^m = \log[-\log(\hat{\alpha}_j^m)] - \log[-\log(\alpha_j^m)]$. Taking the Taylor series expansion of $\log[-\log(\hat{\alpha}_j^m)]$ around α_j^m, we get

$$v_j^m = \frac{\hat{\alpha}_j^m - \alpha_j^m}{\alpha_j^m \log(\alpha_j^m)} + o\left(|\hat{\alpha}_j^m - \alpha_j^m|\right). \tag{16}$$

For large N (therefore large N^m), the term $o(\cdot)$ can be ignored. Hence the variance of v_j^m, conditional on N^m, can be approximated by the variance of the first term on the right-hand side of (16). Therefore,

$$V_j^m = V[v_j^m | N^m] \approx \frac{V[\hat{\alpha}_j^m]}{[\alpha_j^m \log(\alpha_j^m)]^2}. \tag{17}$$

In matrix form, we can express the system of equations (15) as

$$\begin{pmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_T \end{pmatrix} = \begin{pmatrix} \mathbf{W} & \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{W} & \mathbf{0} & \mathbf{1} & \cdots & \mathbf{0} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathbf{W} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{1} \end{pmatrix} \begin{pmatrix} \beta \\ \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_T \end{pmatrix} + \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_T \end{pmatrix}, \tag{18}$$

[9] We will investigate various ways to estimate S_j^m under different data settings in Section 4.3.

or more concisely,

$$\tilde{\mathbf{Y}} = \mathbf{X}\mu + \tilde{\mathbf{V}}. \tag{19}$$

It can be easily shown that $\Omega = Var[\tilde{\mathbf{V}}]$ has a simple heteroskedasticity structure asymptotically and can be consistently estimated. The minimum chi-square estimator of $\mu = (\beta', \lambda_1, ...\lambda_T)'$ is defined as the Aitken generalized least squares estimator. From (19),

$$\hat{\mu} = \left(\mathbf{X}'\hat{\mathbf{\Omega}}^{-1}\mathbf{X}\right)^{-1} \mathbf{X}'\hat{\mathbf{\Omega}}^{-1}\mathbf{Y}. \tag{20}$$

It is well known that the minimum Chi-square estimator has the same first-order asymptotic distributions as the maximum likelihood estimator, provided that the group size N^m goes to infinity as the sample size increases. The proof for the consistency and asymptotic normality of the minimum Chi-square estimator is now standard. See, for example, Amemiya (1980, 1985) for a general treatment and Ryu (1994) for the application to univariate proportional hazard duration models.

The asymptotic variance-covariance matrix Σ_1 of the minimum Chi-square estimator can be easily estimated as,

$$\hat{\Sigma}_1 = \left(\mathbf{X}'\hat{\mathbf{\Omega}}^{-1}\mathbf{X}\right)^{-1}. \tag{21}$$

4.2 A Simple Heteroskedasticity Robust Procedure

The minimum Chi-square estimator defined earlier is in principle very easy to compute. The most complicated aspect is the weight adjustment in the weighted least square step. To explore the seemingly uncorrelated structure, one can simplify the whole procedure by carrying out an ordinary least square (OLS) in the second step,

$$\tilde{\mu} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y}. \tag{22}$$

The efficiency loss using $\tilde{\mu}$ in the current framework can be assessed using Monte Carlo experiments.[10] Of course, one needs to correct the estimation using the

[10] We leave this assessment for future research and only report our experience with the particular empirical application in section 6.

heteroskedasticity-robust variance-covariance estimator of White (1980),

$$\Sigma_o = V[\tilde{\mu}] = (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{\Omega}\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}.$$

To estimate Σ_o, one uses the squared OLS residues to estimate the diagonal matrix Ω.

4.3 Estimation of Cell-Specific Survivor Functions

From the previous discussion the minimum Chi-square estimation of the semiparametric survivor function (3) requires a consistent estimation of S_j^m and its sampling variance for all j and m. We now review various methods available for different data settings.

By assumption the covariates X are weakly exogenous to the model of willingness to pay. Each cell m consists of a homogeneous sub-population. This implies that the estimation of S_j^m is separable across cells. In what follows we suppress the notation for the cell indicator with the understanding that we will follow the same procedure for each cell.

When single-bounded dichotomous choice questions are used in a CV survey, each observation is either left- or right-censored. The "pooling adjacent violators" algorithm of Ayer *et al* (1955) can be applied.

When data consist of only interval- and right-censored observations — a situation not applicable to CV surveys — the Kaplan-Meier product limit estimator can be applied to consistently estimate the survivor function.

When double-bounded dichotomous choice questions with overlapping bids are used in a CV survey,[11] each observation is either right-, interval-, or left-censored.

[11]See An and Ayala (1996) for a detailed discussion of the two types of data grouping mechanisms that can be encountered when using double-bounded dichotomous choice questions: overlapping bids and distinct bids.

For this case, Turnbull (1974) introduced the self-consistent algorithm. The basic logic of Turnbull's algorithm is to use the current trial estimates of the survivor function to reallocate the left-censored spells and then apply the Kaplan-Meier estimator to the adjusted data to derive the new trial estimates of the survivor function. This iterative algorithm converges quickly to a solution which is equivalent to the non-parametric maximum likelihood estimates for the survivor function.

When double-bounded dichotomous choice questions with distinct or mixed bids are used in a CV survey, each observation is either right-, interval-, left- or across-interval-censored. For this case, the generalized self-consistent algorithm proposed by An and Ayala (1996) can be used.

Because the distinct bids double-bounded dichotomous choice questions are the most popular survey methods by far, and because the An and Ayala algorithm applies to all other special cases just mentioned, we now concentrate on a description of it. The generalized self consistent algorithm maintains the same *reallocation* idea introduced by Turnbull.

The iterative procedure goes as follows. Start with data information as in (6). Given a set of starting values $\{S_j^0, 1 \leq j \leq T\}$ that satisfies the restriction $1 = S_0^0 > S_1^0 > S_2^0 > ... > S_T^0 > 0$,

1) Calculate

$$\delta_j = \sum_{I=0}^{j-1} \sum_{k=j}^{T} \gamma_{ik} \frac{S_{j-1}^0 - S_j^0}{S_i^0 - S_k^0}, \qquad 1 \leq j \leq T, \tag{23}$$

which uses the estimated conditional probabilities as weights to reallocate all the across-interval and left-censored observations into an adjusted number of interval-censored observations.

2) Calculate

$$\begin{cases} S_1^1 = 1 - \frac{\delta_1'}{n_1} \\ S_j^1 = \left[1 - \frac{\delta_j}{n_j}\right] S_{j-1}^1, \ j = 2, 3, ..., T, \end{cases} \tag{24}$$

where $n_j = \sum_{k=j}^{T}(\gamma_{k,T+1} + \delta_k)$. Equation (24) would be the Kaplan-Meier estimator if the data consisted of interval- and right-censored observations.

3) Return to step 1 and replace $\{S_j^0, 1 \leq j \leq T\}$ with $\{S_j^1, 1 \leq j \leq T\}$, etc.

4) Stop when the required accuracy has been achieved.

An and Ayala show that the solution of this algorithm is exactly the solution to the distribution free log-likelihood function subject to the monotonicity constraint. Hence, it is $\sqrt{N}$ consistent, where N is the sample size.

5 Other Issues

Estimation of the semiparametric willingness to pay distribution is of course only the first step. This section deals with some related issues.

5.1 Testing the Proportional Hazard Specification

Even though the specification (2) achieves maximum flexibility within the proportional hazard framework, the proportionality is itself an assumption which should be tested.[12]

Many procedures have been proposed to test the proportionality assumption in duration analysis. The test statistics are often derived using generalized residues. This requires observation of actual values of the dependent variables. In the current context where every observation of the willingness to pay is grouped, these generalized residues cannot be calculated. One of the tests, proposed by Ryu (1994), seems most appropriate.

[12]In duration analysis, the proportional hazard model (PH) specification has been criticized for lack of justification. An (1995) provides sufficient and necessary conditions for the PH specification in a controlled jump process framework.

Ryu proposes to estimate the model with artificial further aggregation of the already grouped data. Let the original estimator for β be $\hat{\beta}$ and the new estimator be $\tilde{\beta}$. Ryu's test exploits the fact that under the null hypothesis that the proportionality is correct, both estimators will be consistent. That is, if proportionality holds, both $\tilde{\beta}$ and $\hat{\beta}$ will converge to the same β. Otherwise they will converge to different quantities. The test statistic is then based on the distance of the two estimates,

$$R = (\tilde{\beta} - \hat{\beta})'[V(\tilde{\beta} - \hat{\beta})]^{-1}(\tilde{\beta} - \hat{\beta}). \tag{25}$$

This is a little different from a Hausman test since the two estimators do not necessarily satisfy an informational nesting relation which was the key for a Hausman test. When the original "finer" data are used in the estimation, we are estimating possibly more points of the baseline hazard function. This will not necessarily make $\hat{\beta}$ more efficient than $\tilde{\beta}$. So the estimation of the variance of the difference between the two estimators is a little involved. Ryu (1994) proposes a way to facilitate this calculation.

In the current setting, a natural further integration is for each respondent n, combine $d_n = 2$ and $d_n = 3$ so that the distribution support $[0, I_n]$ will be partitioned into three intervals $[0, B_n^l)$, $[B_n^l, B_n^h)$, and $[B_n^h, I_n]$, instead of the four-interval partition described in Section 2.

5.2 Estimating Mean Willingness to Pay

One of the purposes of estimating the willingness to pay distribution is to assess the welfare measure of the proposed good in question. Two of the most popular measures are the mean willingness to pay, $\mu(X) \equiv E[Y|X]$, and the median willingness to pay, $m(X) \equiv \inf\{y : S(y|X) = 0.5\}$, both of which are conditional on the observed covariates X. Under the proportional hazard specification and with the notation $\psi = \exp\{\beta' X\}$, we have

$$\mu(X) = \int_0^\infty S(y|X)\, dy = \int_0^I e^{-\psi H(y)} dy, \tag{26}$$

and $m(X) = H^{-1}(\psi \log 2)$, where H^{-1} is the inverse function of H.

From the discussion in section 3.2, without parametrization of the base line hazard $h_0(y)$, grouped data from contingent valuation surveys identify at most T number of points of the integrated baseline hazard $H(y)$. Therefore the survivor function of the willingness to pay can be identified at at most T discrete points. To estimate the conditional mean and the conditional median defined above, some interpolation procedure has to be adopted. Three alternatives are available.

(a) One alternative advocated by contingent valuation researchers is the so called conservative estimation which puts all the point mass $S_j - S_{j-1}$ on the left end b_{j-1} of the corresponding interval (Figure 2). In this case,

$$\mu(X) = \sum_{j=1}^{T+1} e^{-\psi\delta_j}(b_j - b_{j-1}) = \sum_{j=1}^{T} e^{-\psi\delta_j}(b_j - b_{j-1}). \tag{27}$$

(b) The second alternative is to take the linear interpolation of the survivor $S(y|X)$. In this case

$$\mu(X) = \frac{1}{2}\sum_{j=1}^{T+1}[e^{-\psi\delta_j} + e^{-\psi\delta_{j-1}}](b_j - b_{j-1}). \tag{28}$$

Under the current setting, it is probably natural to adopt the linear interpolation of the integrated baseline hazard function. In this case, it can be shown that,

$$\mu(X) = \sum_{j=1}^{T} e^{\psi\delta_{j-1}}\left[1 - e^{-\psi(\delta_j - \delta_{j-1})}\right]\frac{b_j - b_{j-1}}{\psi(\delta_j - \delta_{j-1})}. \tag{29}$$

5.3 Interpreting β and Estimating Marginal Effects

One of the main advantages of the semiparametric specification over its nonparametric counterpart is that with the former one can estimate the marginal effect of the social and economic variables on the (mean) willingness to pay. Under the proportional hazard specification (2),

$$\beta = \frac{\partial}{\partial X}\log h(y|X),$$

are the marginal effects of the covariates on the logarithm of the hazard function. More direct quantities in the willingness to pay context are the marginal effects of the covariates on the conditional mean willingness to pay,

$$\kappa = \frac{\partial}{\partial X}\mu(X) = -\beta\psi \int_0^I e^{-\psi H(y)} H(y) dy$$

which can be estimated by choosing one of the three interpolating methods as discussed earlier.

6 An Empirical Example

To illustrate the statistical inference procedures proposed in the previous sections, we use data from a contingent valuation study conducted by Hanemann *et al* (1991) to elicit the WTP for protecting wetland habitats and wildlife in California's San Joaquin Valley.

6.1 The Data

Hanemann *et al* evaluate five different environmental programs in their survey. We focus on the respondents' responses to a wetlands and wildlife improvement program. This survey was conducted via a combination of mail and telephone media. Respondents were chosen based on random digit sampling and were asked whether they were willing to participate in a survey. If they agreed, a mail questionnaire was sent and a certain time was arranged for the interviewer to call the household. The survey used a double-bounded dichotomous choice format. The sets of bids used in this study are: (25, 55, 110), (30, 65, 125), (40, 75, 125), (40, 80, 125), (65, 125, 170), (75, 140, 250), (80, 170, 250), (125, 210, 375) and (125, 250, 375). These 9 versions of the questionnaires involve $T = 14$ distinct bids. The interview process was conducted in May 1989 in three geographical areas: (1) the San Joaquin Valley, (2) the rest of the state of California, and (3) the states of Oregon, Washington and Nevada as representatives of the "out-of-state" population. Of the 1960 households originally contacted, 1239 (63.1%) agreed to participate. From these, 1004 completed interviews were collected.

In this empirical application we use the 576 interviews from the rest of the California sample. After deleting those with no information on income and those who answered "don't know" to the original bid, only 532 observations are used.

In addition to the willingness to pay questions, the data set also records respondents' personal characteristics. Among the covariates, residence in California and age are measured in years; sex, race and membership of an environmental organization are categorical; education level, labor supply and annual income are measured in groups. Table 1 reports the sample summary statistics. Table 2 reports for the whole 532 observation the across-interval-censoring count of the willingness to pay corresponding to the data information matrix (6).

6.2 Results

In order to carry out the minimum Chi-square estimation we have further collapsed the 8 income groups into high income and low income, and reclassified the 8 education levels into with or without education beyond a high school degree. We have divided the whole sample into 8 cells according to sex (female, male), income (low, high) and education (low, high) groupings. For example, a respondent in Cell 1 is a female with low income and without higher education. A respondent in Cell 8 is a male with high income and with higher education. Nonparametric estimations of the survivor functions for the whole sample and for each cell using An-Ayala algorithm are reported in Table 3. Pulling in the sense of Ayer *et al* is common. Notice that for Cells 3 and 7 there are not enough observations to obtain estimates for the survivor probabilities. For the linear regression step in the minimum Chi-square estimation, these two cells will not be used.

Table 4 reports the maximum likelihood estimates for $\theta = (\beta', \lambda_1, ..., \lambda_T)'$. To start our empirical investigation, we first estimated a parametric model using Weibull distribution. Weibull family is a special case of the proportional hazard models with the baseline hazard $h_0(y) = \eta y^{\eta-1}$. A Weibull distribution has a monotone hazard function whose shape is determined by the parameter η. When $\eta = 1$ the Weibull further specializes into an exponential distribution whose hazard is independent of y. The estimated shape parameter, $\hat{\eta} = 1.0476$ is not significantly different from unity, suggesting constant hazard (Panel I).

Panel II in Table 4 reports the parameter estimates using the proportional hazard model specification. Since the proportional hazard specification nests the Weibull model, a likelihood ratio test can be conducted. Such a test reveal that the Weibull specification can not be rejected within the more general proportional hazard specification for the current data set.

Panel III in Table 4 reports the parameter estimates using further aggregated data as described at the end of section 5.1. These estimates are meant to test the proportionality assumption. Due to the special feature of the questionnaire design, for the current setting it turned out that the further aggregated data identify the same number of parameters as the original data, so the informational nesting relation discussed in section 5.1 holds. This makes the calculation of the variance matrix for the difference of the two sets of estimates very easy. Such a test can not reject the proportionality assumption.

Table 5 reports the parameter estimates using both the maximum likelihood method and the minimum Chi-square method. To make the two sets of the estimates comparable, we have used the three binary grouped covariates, female, higher education and higher income. For the minimum Chi-square estimates, we have used the simple OLS procedure in the second step which ignores the heteroskedasticity. The big difference in the estimated standard errors, relative to their maximum likelihood counterparts, demonstrates the efficiency loss of such a procedure. On the other hand, the closeness of the two sets of parameter estimates shows that the minimum Chi-square procedure, while easy to compute, generates reliable estimates.

7 Conclusions

This paper is concerned with the estimation of the willingness to pay distribution with data from contingent valuation surveys where, due to the preference for what is called "close-ended" questions, the respondents' willingness to pay is measured only up to interval censoring. With the conviction that in such an estimation, neither fully parametric models nor non-parametric models are fully satisfactory, this paper first propose to use the proportion hazard model specification to represent the willingness to pay distribution. This semiparametric framework, on the one hand, controls for the observed heterogeneity to help researchers and policy makers identify key social economic quantities that determine individuals' willingness to pay, and, on the other hand, allows for maximum flexibility in the shapes of the willingness to pay distribution so that the two key features found by contingent valuation researchers, i.e., asymmetry and multi-modality, can be accommodated.

Data censoring is common in economic duration analysis where the semiparametric proportional hazard specification is borrowed. However, in duration analysis researchers often observe complete durations, i.e., exact values of the dependent variables, for at least some, if not all, sample points. The rest of the sample points are often assumed to be right censored.[13] Analysis of grouped duration data is relatively new (Kiefer, 1988, Ryu, 1994, Sueyoshi, 1995). Even there the data grouping mechanism is considered very simple. It is quite normal in duration surveys that a duration is known to fall within a natural interval (a month, a year, etc.). In contingent valuation surveys with double bounded dichotomous choice questions, the data grouping mechanism is more complicated. Often "across-interval-censored" data on the dependent variable are observed. Nevertheless the main apparatus for statistical inference remains the same, as this paper demonstrated.

With only across-interval-censored data, there is a maximum flexibility of the

[13]Left censoring can be extremely difficult to deal with. See for example, Heckman and Singer (1986) and An (1996a).

model structure which can be identified. With grouped data, even fully nonparametric specification collapses into essentially finite-dimensional model (Turnbull, 1974). Under the proportional hazard model specification, the maximum flexible structure that can be identified from the data is at most T number of discrete points of the integrated hazard function. This identification result has two important applications. First, the model is essentially finite dimensional and the standard maximum likelihood estimation applies. When the observed covariates in the data fall into the "many-observations-per-cell" type, the minimum Chi-square method is preferred for its ease in computation and programming. This is so thanks to the availability of the An-Ayala (1996) algorithm. Second, as in any nonparametric estimation, the mean willingness to pay and other welfare measures are not identifiable from grouped data. Ad hoc interpolation of either the integrated hazard function or the survivor function has to be adopted. An ad hoc procedure may or may not be conservative, and the approximation error cannot be assessed.

Proportional hazard specification is substantially more flexible than a parametric distribution family. However the proportionality itself is an assumption which can and should be tested. With only across-interval-censored data, feasible testing procedures are limited. This paper advocates the one that uses the further aggregated data from the already grouped data and constructs the test statistic using the difference of the estimated coefficients associated with the observed covariates in the model.

References

Amemiya, Takeshi (1980): "The n^{-2}-Order Mean Square Errors of the Maximum Likelihood Estimator and the Minimum Logit Chi-Square Estimator," *Annals of Statistics*, Vol. **8**, pp. 498-505.

Amemiya, Takeshi (1985): *Advanced Econometrics,*Cambridge: Harvard University Press.

An, Mark Y. (1995):"Econometric Analysis of Sequential Discrete Choice Models," Paper presented at the 7th World Congress of the Econometric Society, Tokyo, Japan.

An, Mark Y. (1996a):"Initial Conditions Problem in Event History Analysis: An Indirect Inference Procedure," in *Life Time Data Models in Reliability and Survival Analysis*, (Jewell, N.P. *et al* eds.) Kluwer Academic Publisher, 7-11.

An, Mark Y. (1996b): "A Proportional Hazard Model for Bivariate Survival Times and Statistical Inference with Grouped Data," manuscript in preparation, Duke University.

An, Mark Y. and Roberto A. Ayala (1995): "A Mixture Model of Willingness to Pay Distributions," Working Paper 95-21, Duke University.

An, Mark Y. and Roberto A. Ayala (1996): "Nonparametric Estimation of a Survivor Function with Across-Interval-Censored Data," Working Paper 96-0?, Duke University.

Ayer, Miriam, H. D. Brunk, G. M. Ewing, W. T. Reid and Edward Silverman (1955): "An Empirical Distribution Function For Sampling with Incomplete Information," *Annals of Mathematical Statistics*, Vol. **26**, pp. 641-647.

Berkson, J. (1953): "Application of the Logistic Function to Bio-Assay," *Journal of the American Statistical Association*, Vol. **39**, pp. 357-365.

Carson, Richard T., W. Michael Hanemann, Raymond J. Kopp, Jon A. Krosnick, Robert C. Mitchell, Stanley Presser, Paul A. Ruud, and V. Kerry Smith (1996), "Was the NOAA Panel Correct about Contingent Valuation?" Discussion Paper # 96-20, Resources for the Future.

Cosslett, S. (1983):"Distribution-Free Maximum Likelihood Estimation of the Binary Choice Model," *Econometrica*, Vol. **51**, pp. 765-781.

Cox, David R. (1972):"Regression Models and Life Tables," *Journal of the Royal Statistical Society, Series B*, Vol. **34**, pp. 187-220.

Cox, David R. and David Oakes (1984): *Survival Analysis*, London: Chapman and Hall.

Cummings, R.G., D.S. Brookshire and W.D. Schulze (1986): *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method,* Totowa, NJ: Rowman & Allanheld.

Han, A. (1987): "Non-parametric Analysis of Generalized Regression Model: The Maximum Rank Correlation Estimator," *Journal of Econometrics*, Vol. **35**, pp. 303-316.

Hanemann, W. Michael (1984): "Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses," *American Journal of Agricultural Economics*, Vol. **66**, pp. 332-341.

Heckman, J.J. and B. Singer (1986):"Econometric Analysis of Longitudinal Data," in *Handbook of Econometrics Vol. III*, (Z. Griliches and M.D. Intriligator eds.), pp. 1917-1978.

Horowitz, Joel L. (1992):"A Smoothed Maximum Score Estimator for the Binary Response Model," *Econometrica*, Vol. **60**, 505-531.

Hausman, Jerry (1993): *'Contingent Valuation: a Critical Assessment,* Amsterdam: North-Holland.

Huang, J-C, Douglas W. Nychka and V. Kerry Smith (1995):"Nonparametric Discrete Choice Methods for Measuring Economic Values," Manuscript, Duke University.

Ichimura, H. (1993):"Semiparametric Least Squares and Weighted Least Squares Estimation of the Single Index Models," *Journal of Econometrics*, Vol. **58**, pp. 71-120.

Kiefer, Nicholas M. (1988): "Analysis of Grouped Duration Data," *Contemporary Mathematics*, Vol. **80**, pp. 107-137.

Klein, Roger W. And Richard H. Spady (1993):"An Efficient Semiparametric Estimator for the Binary Response Models" *Econometrica*, Vol. **61**, pp. 387-421.

Lancaster, Tony (1990): *The Econometric Analysis of Transition Data*, New York: Cambridge University Press.

Manski, Charles (1975):"The Maximum Score Estimator of the Stochastic Utility Model of Choices," *Journal of Econometrics*, **3**, 205-228.

McFadden, Daniel (1994): "Contingent Valuation and Social Choice," *American Journal of Agricultural Economics*, Vol. **76**, pp. 689-708.

Mitchell, Robert C. and Richard T. Carson (1989): *Using Surveys to Value Public Goods: The Contingent Valuation Method,* Washington DC, Resources for the Future.

Ryu, Keunkwan (1994): "Group Duration Analysis of the Proportional Hazard Model: Minimum Chi-squared Estimators and Specification Tests," *Journal of the American Statistical Association*, Vol. **89**, pp. 1386-1395.

Sueyoshi, Glenn (1995):"A Class of Binary Response Models for Grouped Duration Data," *Journal of Applied Econometrics*, Vol. **10**, pp.

Turnbull, Bruce (1974): "Nonparametric Estimation of a Survivorship Function with Doubly Censored Data," *Journal of the American Statistical Association*, Vol. **69**, pp. 169-173.

Wang, W. and M. Zhou: "Iterative Least Squares Estimator of Binary Choice Models: A Semi-parametric Approach," Working Paper, University of Kentucky.

White, Halbert (1980):"A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, Vol. **48**,pp. 817-838.

Zellner, Arnold and T. Lee (1965):"Joint Estimation of Relationships Involving Discrete Random Variables", *Econometrica*, Vol. **33**, pp. 383-394.

Table 1: Sample Summary Statistics

Variable	Mean	Standard Error
Yearca	28.95	16.16
Female	0.5075	0.5004
Age	42.1	15.6
Educ	13.8	2.32
Working	0.9549	0.2077
Envorg	0.1786	0.3834
White	0.7838	0.4120
Income Group:		
Income 1	0	0
Income 2	0.0921	0.2894
Income 3	0.1654	0.3719
Income 4	0.1598	0.3667
Income 5	0.1692	0.3753
Income 6	0.2331	0.4232
Income 7	0.0883	0.2841
Income 8	0.0602	0.2380
No. of Obs.	532	

• Years in residence in CA (Yearca) and Age are measured in years. Education (Educ) is measured in 8 schooling levels. Female, Envorg, Working and White are all binary indicators. Working=1 for part-time workers as well as full-time workers.

Table 2: Summary of the San Joaquin Study Data from a Double-Bounded Dichotomous Choice Setting with Mixed Bids

j	b_j	γ_{j1}	γ_{j2}	γ_{j3}	γ_{j4}	γ_{j5}	γ_{j6}	γ_{j7}	γ_{j8}	γ_{j9}	$\gamma_{j,10}$	$\gamma_{j,11}$	$\gamma_{j,12}$	$\gamma_{j,13}$	$\gamma_{j,14}$	$\gamma_{j,15}$
0	0	5	3	15	0	21	8	13	0	28	0	2	0	2	0	0
1	25	-	0	0	5	0	0	0	0	0	0	0	0	0	0	0
2	30	-	-	0	0	4	0	0	0	0	0	0	0	0	0	0
3	40	-	-	-	0	0	3	11	0	0	0	0	0	0	0	0
4	55	-	-	-	-	0	0	0	12	0	0	0	0	0	0	1
5	65	-	-	-	-	-	0	0	0	26	0	0	0	0	0	1
6	75	-	-	-	-	-	-	0	0	12	10	0	0	0	0	0
7	80	-	-	-	-	-	-	-	0	20	0	19	0	0	0	3
8	110	-	-	-	-	-	-	-	-	0	0	0	0	0	0	10
9	125	-	-	-	-	-	-	-	-	-	0	24	16	5	0	75
10	140	-	-	-	-	-	-	-	-	-	-	0	0	17	0	0
11	170	-	-	-	-	-	-	-	-	-	-	-	0	18	0	50
12	210	-	-	-	-	-	-	-	-	-	-	-	-	0	11	0
13	250	-	-	-	-	-	-	-	-	-	-	-	-	-	5	54
14	375	-	-	-	-	-	-	-	-	-	-	-	-	-	-	23

• γ_{ik} = the number of respondents whose willingness to pay falls between bids b_i and b_k with $0 \le i < k \le T+1$ with the convention that $b_0 = 0$, $b_{T+1} = I$, and $\gamma_{0,T+1} = 0$. For definition see (6) in the text.

Table 3: Nonparametric Estimates of Cell Specific Survivor Probabilities Using An-Ayala Algorithm

j	Overall	Cell 1	Cell 2	Cell 3	Cell 4	Cell 5	Cell 6	Cell 7	Cell 8
0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1	0.9153	1.0000	1.0000	NaN	0.9418	1.0000	0.8472	NaN	0.9281
2	0.9153	0.8881	1.0000	NaN	0.9418	0.9051	0.8472	NaN	0.9281
3	0.8950	0.8267	1.0000	NaN	0.9418	0.7311	0.8472	NaN	0.8561
4	0.8251	0.7762	0.9098	NaN	0.9418	0.7311	0.8336	NaN	0.8561
5	0.8129	0.7762	0.8121	NaN	0.9001	0.7311	0.8232	NaN	0.7845
6	0.8129	0.7762	0.8121	NaN	0.8689	0.7311	0.8232	NaN	0.7845
7	0.7914	0.7690	0.8121	NaN	0.7016	0.7311	0.7860	NaN	0.7567
8	0.6037	0.6886	0.6096	NaN	0.6501	0.4257	0.7860	NaN	0.5821
9	0.6037	0.6083	0.6096	NaN	0.6501	0.4256	0.5494	NaN	0.5821
10	0.6037	0.6083	0.6096	NaN	0.6501	0.3174	0.4001	NaN	0.4992
11	0.4652	0.4951	0.3756	NaN	0.4912	0.3174	0.4001	NaN	0.4578
12	0.3688	0.3918	0.3756	NaN	0.4137	0.2128	0.2888	NaN	0.2911
13	0.3242	0.2836	0.2695	NaN	0.3795	0.1587	0.2888	NaN	0.2851
14	0.2074	0.2836	0.2695	NaN	0.1182	0.0000	0.2311	NaN	0.2425

• Cell 1: female-low income-poor educ; Cell 2: female-low income-high educ; Cell 3: female-high income-poor educ; Cell 4: female-high income-high educ; Cell 5: male-low income-poor educ; Cell 6: male-low income -high educ; Cell 7: male-high income-poor educ; Cell 8: male-high income-high educ.

• For Cells 3 and 7, there are not enough data points to sustain an estimation of the nonparametric survivor functions. The An-Ayala algorithm is described in section 4.3 of the text. A FORTRAN subroutine is available via anonymous ftp on lewis.econ.duke.edu under pub/man/programs subdirectory.

Table 4: **Parameter Estimates from Maximum Likelihood Esimation**

	Weibull Model (I)		**PH with Original Data (II)**		**PH with Further Aggregated Data (III)**	
	Estimate	s. d.	Estimate	s. d.	Estimate	s. d.
Coefficients:Age	**0.0015**	0.0044	**0.0108**	0.0072	**0.0076**	0.0132
Yearca	**-0.0085**	0.0045	**-0.0086**	0.0082	**-0.0084**	0.0149
White	**-0.1953**	0.1357	**-0.0357**	0.9966	**-0.0506**	0.9993
Envorg	**-0.3729**	0.1594	**-0.0539**	1.0033	**-0.0772**	0.9999
Female	**-0.1833**	0.1171	**-0.0108**	0.9966	**-0.0167**	0.9998
Education	**0.0103**	0.0264	**0.0107**	0.1557	**0.0179**	0.5653
log(Income)	**-0.5061**	0.0420	**-0.0732**	0.2481	**-0.0693**	0.8260
δ_1	-		**0.0948**	0.0952	**0.0969**	0.0969
δ_2	-		**0.1972**	0.1404	**0.2013**	0.1425
δ_3	-		**0.3108**	0.1817	**0.3176**	0.1839
δ_4	-		**0.3481**	0.1858	**0.3553**	0.1877
δ_5	-		**0.3940**	0.1918	**0.4018**	0.1933
δ_6	-		**0.4314**	0.1957	**0.4396**	0.1970
δ_7	-		**0.4728**	0.2005	**0.4816**	0.2014
δ_8	-		**0.6915**	0.3104	**0.7192**	0.3116
δ_9	-		**0.8852**	0.3792	**0.9259**	0.3739
δ_{10}	-		**1.1482**	0.4833	**1.2062**	0.4673
δ_{11}	-		**1.3640**	0.5508	**1.4335**	0.5197
δ_{12}	-		**1.6949**	0.6904	**1.7835**	0.6266
δ_{13}	-		**1.8367**	0.7174	**1.9288**	0.6432
δ_{14}	-		**2.2266**	0.8564	**2.3308**	0.7586
Weibull η	**1.0476**	0.0711	-		-	
-Log likelihood	**708.83**		**707.79**		**549.11**	

Table 5: **Parameter Estimates Using MLE and the MCS Method**

	MLE (I)		**MCS (II)**	
	Estimate	s. d.	Estimate	s. d.
Coefficients:				
Female	**-0.242**	0.2289	**0.2138**	0.2544
Higher Education	**0.0306**	0.05188	**0.1945**	0.2421
Higher Income	**-0.292**	0.0686	**-0.3999**	0.2652
δ_1	**0.0921**	0.0934	**0.0906**	0.3563
δ_2	**0.1936**	0.1398	**0.1883**	0.4138
δ_3	**0.3060**	0.1831	**0.2953**	0.6499
δ_4	**0.3433**	0.1879	**0.3302**	0.9340
δ_5	**0.3892**	0.1953	**0.3742**	1.3455
δ_6	**0.4266**	0.2004	**0.4092**	1.7308
δ_7	**0.4681**	0.2067	**0.4497**	2.0122
δ_8	**0.6916**	0.3876	**0.6360**	2.3092
δ_9	**0.8881**	0.4889	**0.8079**	2.4995
δ_{10}	**1.1632**	0.6365	**1.0515**	2.6428
δ_{11}	**1.3919**	0.7228	**1.2508**	2.9229
δ_{12}	**1.7606**	1.0723	**1.5370**	3.1116
δ_{13}	**1.9119**	1.0979	**1.6666**	3.4063
δ_{14}	**2.3471**	1.3176	**2.0147**	3.6499
-Log likelihood	**714.33**		**-**	

Figure 1. Identifiable Baseline Hazard Function

(Not available in this electronic version)

Figure 2. Three Alternative Interpolations Needed to Estimate the Mean Willingness to Pay

(Not available in this electronic version)