Messy Data Modelling in Health Care Contingent Valuation Studies

Maria Ana Odejar, Mandy Ryan,and Kostas Mavromaras
University of Aberdeen

Working Paper
Health Economics Research Unit
University of Aberdeen
January 2004 Draft
(Preliminary draft please don't cite)

Abstract

This study addresses the complexity in modeling contingent valuation surveys that account for sample selection bias, true zeros and non-ignorable unresolved cases or missing responses including "don't knows" and protest responses. Estimation and inference on welfare measure such as mean WTP or health care valuation for demanders, non-demanders as well as counterfactual welfare estimates for these unresolved cases are also provided. Two alternative endogenous switching tobit (EST) models are specified to simultaneously estimate the parameters of the latent willingness to pay (WTP) decision variable and the latent true WTP level. Bayesian techniques are developed using Markov chain Monte Carlo (MCMC) methods data augmentation and Gibbs sampling with Metropolis-hastings for estimating the endogenous switching tobit model. The Bayesian approach presented here is useful even for finite sample size and for models with relatively flat likelihood like sample selection models for which convergence is a problem or even if convergence is achieved correlation of the latent random errors are outside the (-1,1) range. The proposed methodology is applied to a single-bounded dichotomous choice contingent valuation model using British Eurowill data on evaluating cancer health care program. Results in this study reveal that the interview interest scores for the unresolved or missing cases are substantially high and not far from scores of "yes" respondents. The pattern in the values of socio-economic and health related variables shows that these unresolved cases are not missing completely at random. Consequently, dropping them from the analysis is both inefficient and biased since they may actually contain valuable information on the willingness decision process and the true WTP level of respondents. Inclusion of these non-ignorable missing values is not detrimental to the quality of parameter estimates as reflected in the sum of log conditional predictive ordinate (SLCPO) goodness-of-fit criterion and smaller standard deviation of parameter estimates. The classification results in this study agree with Haener and Adamowics (1998) and Groothius and Whitehead 2002) findings and contradict a common practice of considering all "don't know" responses as "no" responses. The positive correlation of the latent random errors may explain why the true WTP levels in DC contingent valuation studies are oftentimes overestimated. The model presented in this paper may also be extended to double-bounded dichotomous choice models with slight modification. Although respondents were reminded to ignore the other two programs when assessing the value of each health care program, the model presented here can easily be extended to allow estimation and inference on correlations of valuations among the three health care programs due to factors such as anchoring effect. This extended model can also test for differences in mean WTP among the three health care programs.

1.0 Introduction

Determining the economic value of public or non-market good such as health care intervention is oftentimes of prime interest to policy makers. Dichotomous choice contingent valuation method (DCCVM) is one popular value elicitation scheme first applied by Bishop and Heberlein (1979) while Hanemann (1984) developed the conceptual and theoretical arguments for using this technique to elicit welfare benefits. Arrow and Solow (1993) together with other panel experts appointed by the US National Oceanic and Atmospheric Administration to assess the validity of the contingency valuation (CV) method, recommended the DC method over the open-ended approach and inclusion of “don’t know” (DK) option. In the close-ended DCCVM a random sample of people from the target population are asked if they would be willing to pay for the provision of some public good with option of “yes” , “no” or “don’t know”. Then they are asked if they would be willing to contribute a specific amount for the public good with the same alternative responses “yes” , “no” or “don’t know”. The bid amount is varied across respondents. For the single-bounded (SBDCCVM), only one question is posed to individuals. For the double-bounded (DBDCCVM) each respondent is presented a sequence of two bids. The second bid is conditional on the response to the first bid. It is higher if the response to the first bid is affirmative and lower if negative. For each bid the individual has the options of “yes”,“no” or “don’t know”. The rationale for the inclusion of the “don’t know” is that a considerable number of respondents would prefer this option (Arrow et al. 1993). Moreover, in contingent valuation surveys without the DK option, some of the “yes” or “no” responses may not be meaningful preferences or may not be true reflection of their preferences. A follow-up question is asked for individuals who opted for a “no” response in an attempt to reveal the reasons behind such a negative response. Those who believe that the programme is of no value to their household or cannot afford to pay for its provision may be considered as “true zeros” and others as protests. The protests and DK responses give rise to missing values. A good survey should be able to control or minimize these unresolved cases. However, no matter how much effort is exerted by the interviewer and the question designer, some missing observations would oftentimes be encountered.

There are two general forms of missingness. Item nonresponse is the messy data type where only a part of the information is missing. The other form of missingness is unit nonresponse in which no data at all is available for some sampling units due to unreturned questionaires or refusal of respondents to participate in the survey. Missing data mechanism can be classified into three general categories (Rubin 1976, Little and Rubin 1987). Unobserved values are said to be missing completely at random (MCAR) if missingness is not dependent on known and unknown variables, that is

$P(M/Z,Z^*)=P(M)$

where M is the indicator variable which is 1 if it is observed and 0 if it is missing, Z are observed variables and Z* are latent or unobserved variables. If missingness depends only on observed variables this mechanism is referred to as missing at random (MAR) and $P(M/Z,Z^*)=P(M/Z)$.

Nonignorable missingness occurs when

$P(M/Z,Z^*)$

cannot be simplified further so that missingness depends on unobserved variables aside from the observed variables. It is possible to statistically test the MCAR assumption against the alternate hypothesis that missingness is MAR (Diggle, Liang and Zeger 1994, Little 1988). Without additional information it is impossible to test the MAR assumption against a nonignorable alternative (Little and Rubin 1987). If it is certain that the variable values are missing completely at random (MCAR) then they may simply be dropped from the analysis.

The cost in terms of efficiency and bias from excluding the information from the unresolved or missing cases would be trivial. For data sets with missing values which are not MCAR (NMCAR), it is imperative to strive to satisfy the MAR assumption by measuring covariates and condition on them (Little and Rubin 1999). The most useful covariates for dealing with nonresponse are those that are predictive of the missing data indicator and those that are predictive of the missing variable(s) (Little and Hongyin 2003). A method for analyzing NMCAR incomplete data with covariates is developed by Little(1996). He introduced pattern-mixture models for stratifying incomplete data by the pattern of missing values. Distinct models are formulated with in each stratum. Model parameters are identified by alternative assumptions about the missing data. Maximum likelihood expectation maximization (EM) and stochastic EM (SEM) algorithms and Bayesian Gibbs sampling methods are implemented to provide parameter estimates. An alternative to the pattern-mixture models are endogenous switching regression models (ESRM) with tobit regime models (Cowles, Carlin and Connett 1996) and Odejar and Fahrmeir (2002) of which the selection models (Heckman 1976 and Amemiya 1984) familiar to economists are special cases. The Cowles, Carlin and Connett model is developed for clinical trials with nonignorable missingness. These models simultaneously model missingness and the variable with missing observations. Heckman (1976) suggested a two-stage procedure for parameter estimation of the sample selection model. However, in applying this estimation method, the researcher should ensure that the variance-covariance matrix should be adjusted for heteroscedasticity and the for the fact that the inverse Mills ratio is unobserved and only estimates are used in the second stage of the estimation procedure. Greene (1981) specified the corrected asymptotic covariance matrix. A more efficient estimation method is full information maximum likelihood (FIML) in which parameter estimates are consistent, asymptotically normal and asymptotically efficient and achieves the Cramer-Rao lower bound. FIML may be implemented using nonlinear optimisation methods such as Newton's algorithm. This method is computationally quite cumbersome especially with increasing number of regressors. and it may converge to a local maximum or even to a saddle point. Sometimes convergence values of the random errors correlation are outside the (-1,1) range. Copas(1990) recommended evaluating the likelihood profile $L(\rho \mid \hat{\gamma}, \hat{\beta}, \hat{\sigma}_{\varepsilon})$ for a grid of values of ρ in the (-1,1) interval because the likelihood is well behaved for fixed value of ρ. However, as shown by Copas(1990) and Calia and Strazzera(2001) maximum likelihood estimation methods is not really appealing for sample selection models since these models are often characterized by very flat likelihood functions consequently resulting to convergence problem. Another well known criticism of the maximum likelihood is that it does not provide parameter estimates accurate enough to be useful for small and even moderately large samples. Nawata and McAleer(2002) demonstrated that the finite sample problem with the t-test is alarming and more severe for binary choice and sample selection models. Charlier, Melenberg and van Soest (2001) introduced semiparametric method of estimating the ESRM. Cowles et. al. (1996) and Odejar and Fahrmeir (2002) implemented Markov chain Monte Carlo methods MCMC which provide consistent parameter estimates that are reliable even for finite samples and with convergence easily achieved after a certain burn-in period. Missing responses may also be generated using MCMC data augmentation algorithm. Other imputation methods for contingent valuation studies are discussed in Messonier et. al. (2000). Horowitz and Manski (1998) recommended means for bounding imputations. Brox et. al. (2003) imputed missing WTP responses using Bhat's maximum likelihood based method for categorical data.

In some previous health and environmental contingent valuation studies, the "don't know" responses and protests are dropped from the analysis. A serious cost to this practice is loss of information unless these unresolved cases are similar to the retained sample at least in terms of the socio-economic and health characteristics considered in the study. McClelland and Whittington (1994) however found that DK respondents were more likely female and / or with less education. Haener and Adamowicz (1998) found that DK responses differ in characteristics from the rest of the sample so that deleting them would bias welfare estimate. Jorgensen and Syme (2000) study showed that censoring protest responses would bias CV samples toward those with higher income households.

In the past, protest responses have been incorporated in the non-participation part of sample selection models (Strazzera, et al. 2003, Calia and Strazzera 2001, Alvarez-Farizo et al. 1999 and Donaldson et al. 1998). However, mean WTP is computed only for those who chose to participate in the value elicitation method adjusted for sample selection. No measure of public good value for protest responses are provided. Although those who protest do not conform with the payment vehicle in the survey, the health care program may also be of value to them and they can afford to pay some price unknown to the interviewer.

Analysis of “don’t know” responses have been performed using ordered logistic regression model with the “don’t know” assumed to be a middle response (Groothius and Whitehead 2002) which is a very strong assumption which may be realistic only for ambivalent response but not for those who just cannot make a decision at the time of the interview. In the WTP study of Groothius and Whitehead (2002), the “don’t know” responses are more similar to the “no” responses for their North Carolina sample and more similar to “yes” responses for their Pennsylvania sample. Multinomial model have also offered another method for dealing with “don’t know” responses. However, mean WTP is evaluated indirectly thru the price proxy coefficient. Wang (1997) developed a maximum likelihood procedure that incorporates DK responses in the estimation of WTP assuming that each individual has his/her own implicit valuation distribution and not just a single true value in his/her mind. If the bid is not clearly different from the mean value of one’s own distribution, a DK response is given. This procedure assumes that DK responses are truly uncertain of their preferences at the time of the interview and ignores the possibility of scenario rejection or strategic bias (Haener and Adamowics 1998). Haener and Adamowics (1998) recoded the DK responses to either “yes or “no” based on their reponses to an open-ended payment question. They were asked to indicate how much they are willing to give up annually in exchange of the public good which is old growth forest protection program. If the DK respondents, were willing to give up amounts exceeding the tax level that they were requested to vote on, it is assumed that these DK respondents should have responded “yes” to the referendum question. However, since intervals in the options are quite broad, it is not possible sometimes to determine if the respondents are consistent or not. Of the 67 DK who gave specific answers to the quasi-open ended payment question, 7 were recoded as “yes”, 55 as “no”, 5 are inconsistent and not recoded. Of the 97 DK responses 21 still were unresolved and replied DK to the additional open-ended payment question.

This study develops a Bayesian technique for estimating true willingness to pay or true value of health care program taking into consideration true zeros and nonrandom unresolved cases such as protests, “don’t know” and missing responses. It is assumed that for all respondents including the protests and the “don’t know” respondents, there is some unknown true WTP or true valuation amount which is either above or below the bid given. However, due to some reasons either lack of effort on the part of the question designer, the interviewer or the respondent, the true WTP level may not be elicited giving rise to unresolved cases. In this study, the true level for the unresolved cases is determined to be either above or below the bid using a decision process model which is a function of the individual’s socio-economic and health characteristics decision. The latent true WTP or valuation level is then determined as a function of socioeconomic individual covariates adjusted for the latent decision variable value or net benefit score and truncated to a nonpositive range or a range either above or below the bid. This is a great improvement from the usual method which depends only on the bid level. The MCMC methods data augmentation and Gibbs sampler with metropolis-hastings are implemented for estimating two endogenous switching tobit models in which the true WTP equation is incidentally truncated conditional on the latent decision variable or net benefit score in a single-bounded dichotomous choice contingency valuation exercise. The decision indicator variable for the nonrandom unresolved cases, the values of the latent decision variable or net benefit score, net-benefit threshold value and the latent true WTP levels are considered additional parameters estimated or imputed in the data augmentation step of the MCMC algorithm and are estimated at each iteration step together with the coefficients of the decision process and the true WTP models and the correlation between the unobserved errors in these equations which are generated at the posterior step.

This paper is organized as follows. The next section describes the survey and questionaire design. In section 3, the endogenous switching tobit model and MCMC estimation method are presented. Results of the survey and MCMC simulation procedure are discussed in section 4. Section 5 contains a summary and conclusion of this study.

2. Survey Data

Thirteen Health Promotion Assistants employed by Grampian Health Board, in the Northeastern Scotland conducted questionnaire-based interviews, on behalf of the University of Aberdeen. They sampled the Grampian population in street interviews, various community groups and work environments. Attention was given to obtaining a representative sample of the Grampian population by means of comparison with the most recent census data. This was an ongoing procedure with comparison of demographics being periodically reviewed and if necessary adjustment made to the target population. Respondents were told that taking part in the study would enter them into a prize draw with two prizes of £100 cash.

The questionaire was developed centrally by the Eurowill group (Donaldson, 1999) and customised for the specific country. The questionnaire was concerned with WTP for expansion in 3 health care programs: cancer treatments, heart operations and improved ambulance service for the Grampian population. Respondents were provided with a background information on the type of care currently provided the size of the proposed expansion and the expected health outcomes with and without the expansion. Before evaluating the value of each of the three health care programmes, the interviewer reminds the respondents to ignore the other two programmes for the moment. The initial stage of the WTP exercise is designed to reveal general WTP particpation decision of subjects by asking if they are willing to contribute anything in terms of extra taxation or donation with possible options of “yes”, “no”, and “don’t know”. A follow-up question is posed to determine the motivations behind refusal to pay. The alternative responses are “the programme is of no value to my household”, “other programmes are more valuable”, “other public sector budgets should be cut”, “other groups in society should pay”, “users should pay”, “the health service should be more efficient”, “I can’t afford it”, “I prefer other ways of paying, like private insurance”, “other reasons specified by the respondents”. The second stage question is intended to determine the WTP level of the respondents by asking if their household would be willing to contribute a specific amount annually for the expansion in the number of cancer treatments. Those with affirmative response were asked to explain why they would be willing to contribute.

Table 1 presents the health related and socio-economic variables considered in this study. The interviewer asked if the respondent or anyone in the family had cancer experience. Respondents also rated their health status from 1 indicating very good to 5 indicating very poor. The socio-economic variables are respondent’s age, midpoint of the income category of the household total gross income, number of children under 16 years in the household, respondent's highest educational attainment. The interview related variables are interview interest of the respondent or seriousness to the survey on a 1-5 scale where 1 means not at all interested and 5 means extremely interested, and interview time in minutes.

3. Econometric Framework

Endogenous Switching Tobit Model

This study proposes an endogenous switching tobit (EST) model or incidentally truncated model which simultaneously considers in the estimation procedure both the non-random WTP decision process and the true WTP level which are correlated through the unobserved or latent random errors. The first equation of the EST models the self-selection or the decision process of whether or not to pay for a public good or health intervention. Individuals either deliberately or unconsciously take into consideration the net benefit in terms of expected utility and cost to him of supporting a public good such as health care programme. If the net benefit, $d_i{*}$ is positive, he decides to pay for the cancer health care program either as additional tax or voluntary donation thus accepts the bid d_i is then set to 1. *O*therwise, he refuses to pay and $d_i = 0$.

Self-selection or WTP Decision Model: (1)

$$d_i{*} = z_i'\gamma + \eta_i, \quad i = 1,\ldots,n$$

$$d_i = 1, \quad \text{for } i \in R_1, \qquad \text{if } d_i{*} > 0$$

$$= 0, \quad \text{for } i \in R_{21} \text{ or } R_{22}, \qquad \text{if } d_i{*} \leq 0$$

where d_i* is the latent decision variable for respondent i, d_i is the observed decision dummy variable for the ith respondent, z_i is the vector of socio-economic and health related covariates and η_i are the unobserved random errors in the WTP decision model for the ith respondent.

The second equation of EST is the true WTP level model. Conditional on the latent or unobserved net benefit score, d_i* being positive, the respondent accepts the bid which implies that the latent true WTP amount or reservation price for the health care program is greater than or equal to the bid. When the net benefit d_i* is non-positive, the true WTP or value level is either non-positive which is the case for the true zeros or it is positive but less than the bid so that the bid is refused.

WTP Amount Model: (2)

Regime 1: $w_i^* = x_i'\beta + \varepsilon_i \geq b, \; for \; i \in R_1 \,, \; if \; d_i = 1$

Regime 2:
$$\left. \begin{array}{l} w_i^* = 0 < x_i'\beta + \varepsilon_i < b \,, \quad for \; i \in R_{21} \\ \quad\;\; = 0 \,, \quad for \; i \in R_{22} \end{array} \right\} \quad if \; d_i = 0$$

where γ is the vector of parameters in the WTP decision model, w_i* is the latent true WTP level of the ith individual assumed to follow a normal distribution, x_i is the vector of socio-economic covariates of the latent true WTP level, β is the vector of parameters in the true WTP level model and ε_i are the unobserved random errors in the true WTP level model, and b is the bid . The model specification in (1) and (2) disentangles the confounded direct effects of the covariates and additional effects due to the correlation between the latent random errors in (1) and (2).

The matrix form specification of this model is

$$\boldsymbol{Y} = \boldsymbol{X\Psi} + \xi \qquad (3)$$

where

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_1 \\ \vdots \\ \boldsymbol{Y}_i \\ \vdots \\ \boldsymbol{Y}_n \end{bmatrix} \quad , \quad \boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1 \\ \vdots \\ \boldsymbol{X}_i \\ \vdots \\ \boldsymbol{X}_n \end{bmatrix} \quad , \quad \xi = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_i \\ \vdots \\ \xi_n \end{bmatrix}$$

$$Y_i = \begin{bmatrix} d_i * \\ w_i * \end{bmatrix}, X_i = \begin{bmatrix} z_i' & \\ & x_i' \end{bmatrix}, \boldsymbol{\psi} = \begin{bmatrix} \boldsymbol{\gamma} \\ \boldsymbol{\beta} \end{bmatrix}$$

$$\boldsymbol{\xi_i} = \begin{bmatrix} \boldsymbol{\eta_i} \\ \boldsymbol{\varepsilon_i} \end{bmatrix} \sim N\,(\boldsymbol{0}, \boldsymbol{\Sigma}_\xi) \ , \ \boldsymbol{\Sigma}_\xi = \begin{bmatrix} 1 & \rho_{\eta\varepsilon} \\ \rho_{\eta\varepsilon} & 1 \end{bmatrix}.$$

Since both the latent WTP decision variable and the latent true WTP level or true degree of WTP are unobserved, the coefficients are estimable only up to a scale factor. Thus, both variances of η_i and ε_i are assumed equal to one. The possible responses that are in the first regime are $R_1 = \{I_T^Y I_B^Y, I_T^N I_D^Y I_B^Y, I_T^N I_D^{DK} I_B^Y, I_T^{DK} I_D^Y I_B^Y, I_T^{DK} I_D^{DK} I_B^Y\}$ where I_k^l , k=T,D,B which denote extra tax, voluntary donation and bid respectively, $l \in \{Y,N,DK\}$ indicates the response for the kth stage question. The superscripts *Y, N, DK and M* indicate "yes", "no", "don't know" and missing responses respectively. Those responses that fall in the second regime are $R_{21} = \{I_T^Y I_B^N, I_T^N I_D^Y I_B^N, I_T^N I_D^{DK} I_B^N, I_T^{DK} I_D^Y I_B^N, I_T^{DK} I_D^{DK} I_B^N\}$. True zero WTP level associated with those who consider the program of no value to their household or their household cannot afford to pay belong to R_{22}. Protests and "don't knows" R_M = $\{I_T^Y I_B^{DK}, I_T^N I_D^Y I_B^{DK}, I_T^N I_D^{DK} I_B^{DK}, I_T^{DK} I_D^Y I_B^{DK}, I_T^{DK} I_D^{DK} I_B^{DK}\}$ have missing responses to WTP level question. These missing responses are non-ignorable since the probability that they are missing may actually depend on their unobserved or latent true WTP. Since the protests are the effect of elicitation method rather than a measure of the true valuation of the programs, these responses and the "don't knows" were imputed using socio-economic and health covariate values. The indicator variable d_i for the missing responses are considered additional parameters to be estimated using individual's socio-economic and health related covariate values.

Bayesian Estimation Via Markov Chain Monte Carlo Method

The likelihood of the endogenous switching tobit model in (1) and (2) is

$$L(\theta/Y)=\prod_{i\in R_1} P\left[d^*>0,w^*\geq b\right] \prod_{i\in R_{21}} P\left[d^*\leq 0,0<w^*<b\right] \prod_{i\in R_{22}} P\left[d^*\leq 0,w^*=0\right] \quad (4)$$

where $\boldsymbol{\theta} = \{ \boldsymbol{\gamma}, \boldsymbol{\beta}, \sigma_{\eta\varepsilon} \}$. The detailed version of this likelihood is in appendix A. This likelihood is analytically intractable and is relatively flat Copas (1990) and Calia and Strazzera (2001). Consequently, convergence is difficult to achieve using maximum likelihood algorithms and even if convergence is achieved the correlation may be outside the interval (-1,1).

Bayesian approach greatly simplifies analysis of this endogenous switching tobit model. The key to analysing this model is to apply data augmentation algorithm to generate the missing or latent values $\boldsymbol{M}_y^{(m)} = [\, d_i^M,\, d_i^*\,, w^*]$ and analyse it as a seemingly unrelated regression model in the posterior step. With the original data and these generated latent values, the data is complete and the likelihood simplifies to a multivariate linear model

$$\boldsymbol{L(\theta,Y)} \propto exp\left[-\frac{1}{2}\boldsymbol{tr}\sum_{i=1}^{n}\boldsymbol{\Sigma}_{\xi}^{-1}\left(\boldsymbol{Y}_i-\boldsymbol{X}_i\psi\right)\left(\boldsymbol{Y}_i-\boldsymbol{X}_i\psi\right)'\right]. \quad (5)$$

To ensure a proper posterior density, parameters may be modelled with informative priors. It is trivial to obtain prior information from a subset of the sample data using FIML estimates or probit estimates, cluster analysis and discriminant analysis and even from correlations of the covariates and the dummy variables. MCMC method estimates especially the coefficients of both the decision and the true WTP level equations of the EST are fairly robust to prior information (Odejar and Fahrmeir 2002) and starting values unlike estimates generated by the EM algorithm. The joint prior density of the parameters is

$$\boldsymbol{g(\theta)=g(\Psi,\Sigma_{\xi})=g(\Psi)g(\Sigma_{\xi})} \quad (6)$$

where $g(\boldsymbol{\Psi})$ is a multivariate normal $N(\boldsymbol{A}, \boldsymbol{V}_{\boldsymbol{\Psi}})$ and $g(\boldsymbol{\Sigma}_{\xi})$ is inverse-Wishart so that $\boldsymbol{\Sigma}_{\xi}^{-1} \sim \boldsymbol{W}\left(\alpha, (\alpha\boldsymbol{R})^{-1}\right)$ with mean $\boldsymbol{R}^{-1}$ and precision matrix $(\alpha\boldsymbol{R})^{-1}$. Combining (5) and (6), yields the posterior distribution $g(\boldsymbol{\theta} \mid \boldsymbol{Y})$. MCMC methods contruct a Markov chain $\theta_{\boldsymbol{k}}^{(1)}, \theta_{\boldsymbol{k}}^{(2)}, \ldots, \theta_{\boldsymbol{k}}^{(m)}, \ldots, \theta_{\boldsymbol{k}}^{(t)}$ with equilibrium distribution identical to the desired joint posterior distribution. Ergodic averaging of the Markov chain $\theta_{\boldsymbol{k}}^{(m)}$ or some function $\boldsymbol{h}\left(\theta_{\boldsymbol{k}}^{(m)}\right)$ provides consistent estimators of the parameters $\boldsymbol{\theta}$ or its function $\boldsymbol{h}(\theta)$. A form of MCMC is Gibbs(Geman and Geman 1984) sampling where Markov chains are generated from full conditional distributions. Data augmentation is Gibbs sampling with two blocks. In the imputation step, the missing or latent variables are generated sequentially and in the posterior step, the rest of the parameters are drawn from the conditional distributions. The details of the MCMC algorithm for estimating the parameters of EST in (1) and (2) is in appendix B.The iterations from the imputation and posterior steps of the MCMC method provide a Markov chain with transition probability from $\boldsymbol{\theta}^{(m)}$ to $\boldsymbol{\theta}^{(m+1)}$ given by the product of the full conditional probabilities. Under regulatory conditions (Tierney 1991), as the number of iterations m approaches infinity, $\boldsymbol{\theta}^{(m)}$ converges in distribution to $\boldsymbol{\theta}$. After equilibrium is reached at iteration a, the generated sample values are averaged to provide consistent estimates of the parameters or their function,

$$\hat{\boldsymbol{E}}[\boldsymbol{h}(\boldsymbol{\theta}_{\boldsymbol{k}})] = \sum_{m=a+1}^{t} \frac{\boldsymbol{h}\left(\boldsymbol{\theta}_{\boldsymbol{k}}^{(m)}\right)}{\boldsymbol{t} - \boldsymbol{a}} \tag{7}$$

and the estimate of the conditional predictive ordinate(CPO) is

$$\int \boldsymbol{f}\left(\boldsymbol{Y}^{f} \mid \boldsymbol{\theta}, \boldsymbol{X}^{f}\right) \boldsymbol{g}(\boldsymbol{\theta} \mid \boldsymbol{Y}, \boldsymbol{X}) \boldsymbol{d}(\boldsymbol{\theta}) \approx \frac{\sum_{m=a+1}^{t} \boldsymbol{f}\left(\boldsymbol{Y}^{f} \mid \boldsymbol{\theta}^{(m)}, \boldsymbol{X}^{f}\right)}{\boldsymbol{t} - \boldsymbol{a}} \tag{8}$$

As depicted in the autocorrelation values of the generated parameter estimates in figure 1, the steady state is quickly reached after only less than or about 50 iterations. The MCMC algorithm is implemented with a burn in or warm up period of $a=200$ iterations and monitoring period of $t=5000$ iterations to provide the ergodic average in (7) and (8) or estimate of the posterior mean and CPO. Parameter estimates are considered significant at the 5%(1%) significance level if the interval between the 2.5th (0.5th) and 97.5th (99.5th) percentiles of the MCMC samples from the posterior distribution exclude zero. Model selection is based on the sum of log CPO's (SLCPO) for cross validation sample of size 12. Model goodness of fit is directly proportional to the SLCPO.

Alternative Endogenous Switching Tobit Model

Another model is based on the assumption that those who refuse the bid have a positive net-benefit less than some threshold value d_{th} so that the decision equation becomes

Self-selection or WTP Decision Model: (9)

$$d_i^* = z_i'\gamma + \eta_i \ , \quad i=1,\ldots,n$$

$$d_i = 2, \quad \text{for } i \in R_1, \quad \text{if } d_i^* > d_{th}$$

$$d_i = 1, \quad \text{for } i \in R_{21}, \quad \text{if } 0 < d_i^* < d_{th}$$

$$d_i = 0, \quad \text{for } i \in R_{22}, \quad \text{if } d_i^* < 0 \ .$$

The true WTP model is similar to the previous model in (2) except that there are three regimes now

WTP Amount Model: (10)

Regime 1: $w_i^* = x_i'\boldsymbol{\beta} + \varepsilon_i > b \qquad \text{for } i \in R_1, \quad \text{if } d_i=2$

Regime 2: $w_i^* = 0 < x_i'\boldsymbol{\beta} + \varepsilon_i < b \qquad \text{for } i \in R_{21}, \quad \text{if } d_i=1$

Regime 3: $w_i^* = 0 \qquad \text{for } i \in R_{22}, \quad \text{if } d_i=0$.

The MCMC algorithm for the EST model 2 is similar to the algorithm for the EST model 1 in (1) and (2) except for the steps in generating the threshold value of the net-benefit and the latent net-benefit scores d_i^* for i in $\in R_{21}$ and R_1. The threshold value and latent net-benefit scores are generated together in a block using metropolis-hastings as developed by Cowles (1996) to accelerate convergence and improve the Markov chains mixing properties .

4. Discussion of Results

Of the 342 observations, only 318 are available for analysis, the rest have missing values for the regressor variables. For the estimation part of the analysis, 306 observations are available and 18 observations are used as cross-validation sample. As shown in figure 1.0, for the WTP question, the "yes" response to the bid is 207/306=67.65%. The "no" response is 52/306=17.00% and don't know is 47/306=15.36%. Of those with "no" response to the WTP decision question, the true zeros is 15/306=4.90% and the protests is 21/306=6.86%. True missing values comprises 3/306=1% of the sample. They have missing values for the bid question either due to recording error or error on the part of the interviewer, although they responded "yes" to the tax or donation question. If only the true zeros and "yes" responses are included in the analysis, (21+47+3)/306=23.21% unresolved cases will be dropped from the analysis. The definition of variables considered in the analysis are presented in table 1. Summary statistics in table 2.0 shows that those with "don't know" and protest responses have very similar mean interview interest scores with those "yes" responses. Those with missing responses have slightly lower average interview interest scores.

The pattern in the values of socio-economic and health related variables as reflected in table 2.0 shows that these unresolved cases are not missing completely at random. The cancer experience, education and income average values are obviously lower for these unresolved cases than those with "yes" responses. So that these unresolved cases actually contain valuable information at least on the WTP decision process of respondents.

Parameter estimates provided by a combination of MCMC methods data augmentation and Gibbs sampling are presented in tables 3 for the log-normal form of the endogenous switching tobit model 1. Notice that inclusion of imputed values for the unresolved cases does not sacrifice the quality of parameter estimation as gauged by SLCPO which are quite close. In fact for most of the models considered in this study, the SLCPOs are even higher when unresolved cases are included implying better goodness-of-fit. The parameter estimates of the coefficients and error correlation are similar with and without the unresolved cases. The standard deviations of the parameter estimates are even lower with unresolved cases included. The significant determinants of WTP decision process of respondents are cancer experience, health status, interaction of health and age, income and its square and bid. Note that since the model is in log-form the natural logarithm of all variables except the dummy variables and categorical variables are used in the analysis. Those with cancer experience tend to consider the health care program of greater utility and thus higher net benefit are more likely to decide for supporting the cancer treatment expansion. The cancer health care program is of greater net benefit to the respondents who consider themselves less healthy (that is with higher health score). The health and age interaction has a negative effect on the WTP decision process. Those with higher income, are more likely to decide in favor of supporting the health care program which is consistent with previous studies in the literature. Net benefit of cancer health care program is more elastic to changes in income than its square. Respondents offered lower bids are more likely to decide towards supporting the cancer health care program. For the true WTP level on the other hand, education, income and its square and age and its square have significant effects. True WTP level increases with education level, income, and age. True WTP amount is more elastic to income and age than their squared values.

Evidence of self-selection is reflected in the highly significant correlation between the latent random errors of the WTP decision and WTP level models. Thus, respondents decision on their true WTP level is based on some marginal cost benefit or utility maximizing criterion and not a random process so that it is essential to correct for sample selection bias. The positive correlation of the latent random errors may also explain why the true WTP levels in DC contingent valuation studies are overestimated.

The mean WTP for demanders, non-demanders, true zeros, protests, DKs and missing respondents are shown in table 4. From table 5, it is apparent that although most of the unresolved cases are classified as “no” responses, not all of them can be considered as “no” responses. For protests responses 32.78% were classified as “yes” responses. For the DKs, 30.83% were considered as “yes” responses. For the missing WTP responses, 37% were classifies as “yes” responses.

As depicted in figure 3, convergence is quickly achieved for the coefficients of the decision and the true WTP models and error-correlation for the alternative endogenous switching tobit model 2. However, convergence for the threshold value is very slow. Despite this poor mixing property for the threshold value, parameter estimates for this model which are presented in table 6 are similar in terms of magnitude and signs of the coefficients to those for the EST model 1. However, the error correlations are lower than those for EST model 1 especially when the unresolved cases are excluded. The goodness of fit criterion SLCPO is a little bit higher than those for the EST model 1. Mean WTP in table 7 are also similar to those for EST model 1. The similarity of results maybe explained by the low threshold value which is less than one for both cases when the unresolved cases are excluded and included in the estimation. As expected the remarkable difference is in the number of the unresolved cases classified which are shown in table 8. The percentage of responses classified as yes are 71.74%, 74% and 81.67% for the protest, DKs and missing responses respectively. To policy makers this is good news because these results imply more potential supporters of the health care program.

EST model 2 results agree with those of EST model 1 that inclusion of unresolved cases is not detrimental to the quality of parameter estimates as shown by slightly higher goodness of fit criterion SLCPO and parameter estimates standard deviations which are either lower when unresolved cases are included or close to those when they are excluded. Evidence of sample selection is even stronger as shown by a higher error correlation of 0.86 when unresolved cases are included and only 0.65 when they are excluded.

5. Summary and Conclusion

This study provides a method of estimation and inference about the mean WTP or health care value for demanders, non-demanders, true zeros, as well as protests, DK and missing responses. A Bayesian technique using MCMC methods data augmentation and Gibbs sampling is developed for estimating an endogenous switching tobit model that corrects for sample selection bias and accounts for true zeros and non-ignorable missing responses including “don’t knows” and protests in a single-bounded contingency valuation survey . The Bayesian approach presented here is useful even for finite sample size and for models with relatively flat likelihood like sample selection models for which convergence is a problem or even if convergence is achieved correlation of the latent random errors are outside the (-1,1) range. Results in this study reveal that the interview interest scores for the unresolved or missing cases are substantially high and not far from scores of “yes” respondents. The pattern in the individual specific socio-economic and health characteristics, reveal that these unresolved cases are not MCAR and deleting them from the analysis is inefficient and would bias estimation and inference. Inclusion of these non-ignorable missing values does not jeopardize estimation and the quality of parameter estimates as reflected in the SLCPO goodness-of-fit criterion and smaller standard deviation of parameter estimates. This is true for both EST models 1 and 2. The Bayesian methodology presented in this paper may also be applied to double-bounded dichotomous choice models with slight modification. Moreover, although respondents were reminded to ignore the other two programs when assessing the value of each health care program, the model presented here can easily be extended to allow estimation and inference on correlations of valuations among the three health care programs due to factors such as anchoring effect. This extended model can also test for differences in mean WTP among the three health care programs.

Acknowledgement

The authors gratefully appreciate the data provided by the EUROWILL Group particularly the Aberdeen based study especially Cam Donaldson and David Scott. We also acknowledge the financial support from the Health Foundation, University of Aberdeen and the Chief Scientist Office of the Scottish Executive Health Department.

Figure 1.0

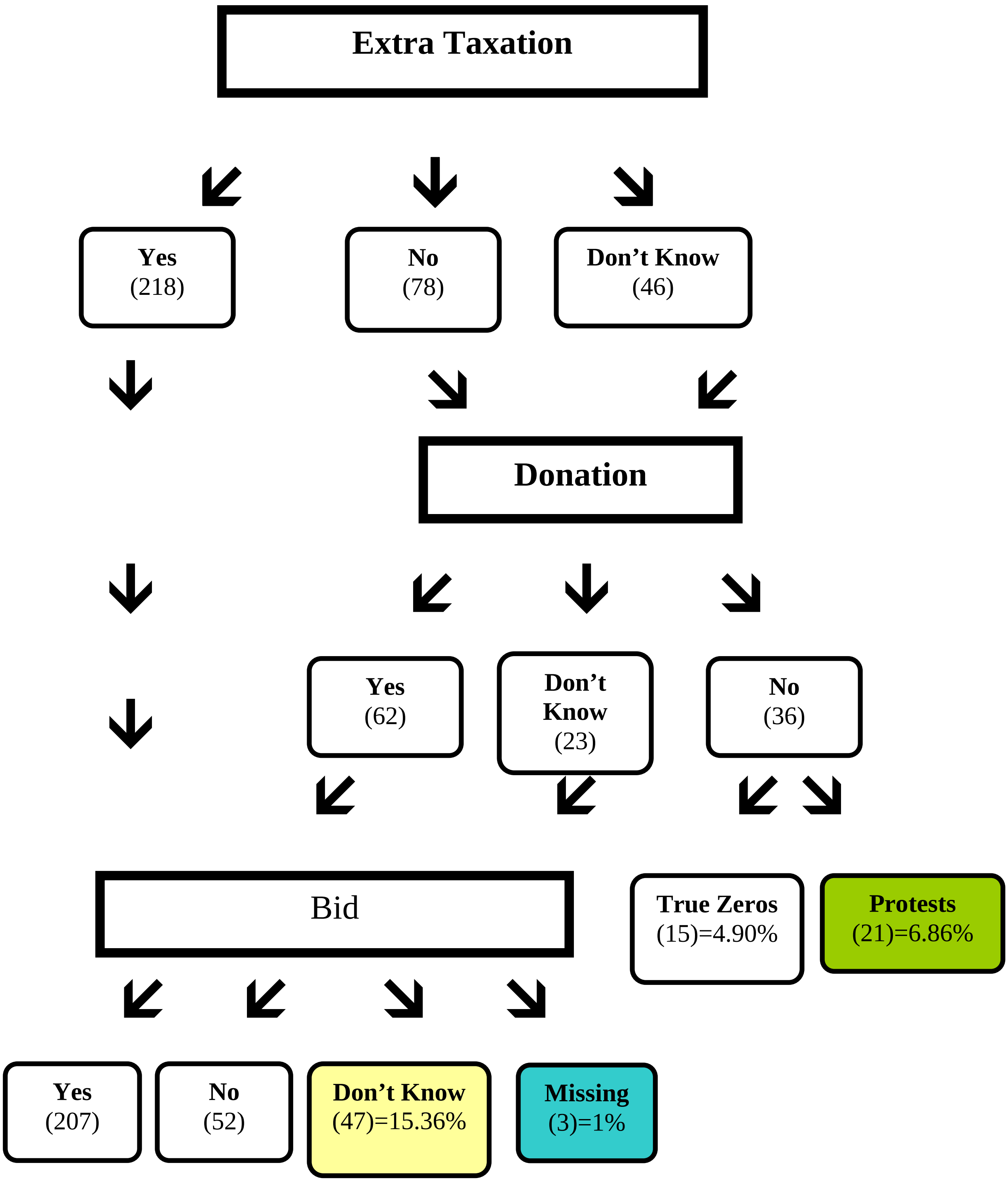
Extra Taxation
Yes
(218)
No
(78)
Don't Know
(46)
Donation
Yes
(62)
Don't
Know
(23)
No
(36)
Bid
True Zeros
(15)=4.90%
Protests
(21)=6.86%
Yes
(207)
No
(52)
Don't Know
(47)=15.36%
Missing
(3)=1%

Figure 2.0 Autocorrelation Functions of the MCMC Estimates

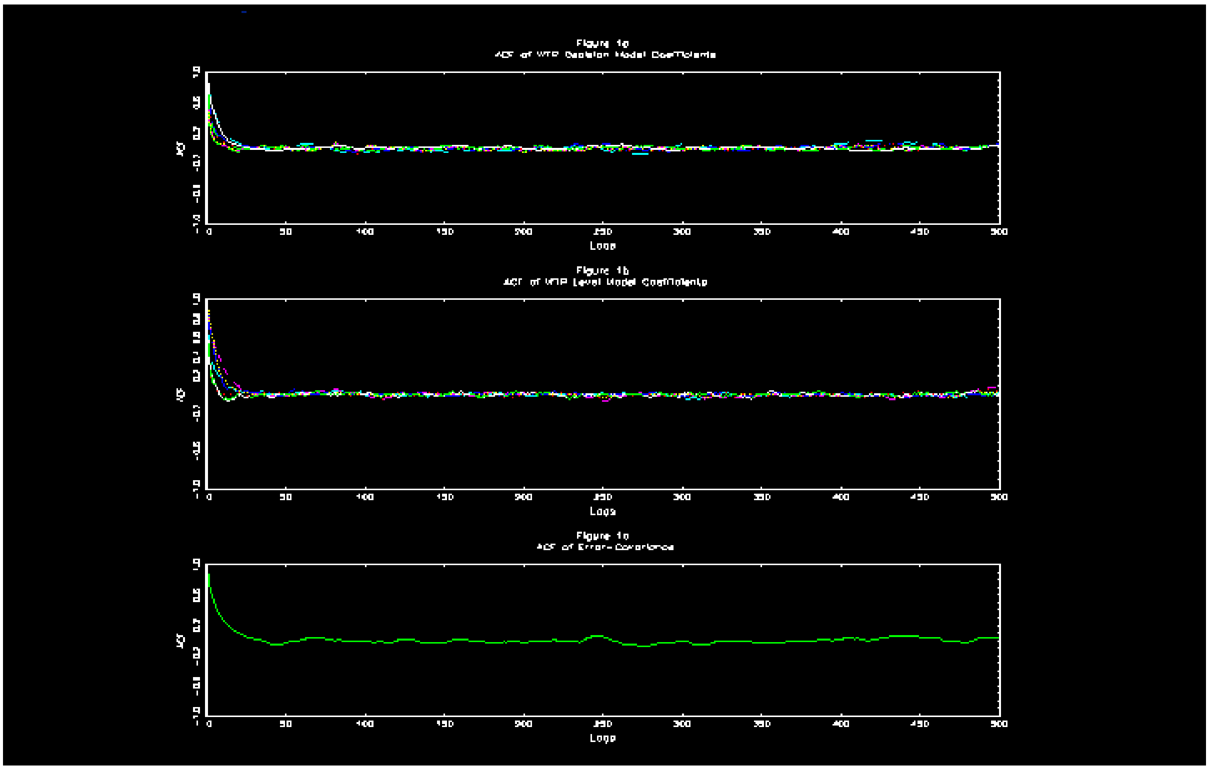

Table 1.0
Description of Variables

Variable Name	Description
Cancer Experience	1= Household has cancer experience 0= No cancer experience
Health Status	1=Very good 2=Good 3=Neither good nor bad 4=Poor 5=Very poor
Age	Respondent's age
Income	Midpoint of household income level
Number of Children	Number of children under 16 years in the household
Education	1=Primary/part secondary 2=Higher level (or A level) 3=Further education college 4=University/polytechnic 6=Other, Specified by respondent
Interview Interest	1-5 scale 1=not at all interested 5 = extremely interested
Interview Time	Length of inteview in minutes

Table 2.0
Summary Statistics:
Mean and
(Standard Deviation)
of Health, Socio-Economic
and Interview Variables

VARIABLES	YES	NO	TRUE ZERO	PROTESTS	DON'T KNOW	MISSING
Cancer Experience	1.35 (0.609)	0.53 (0.513)	0.47 (0.516)	0.63 (0.500)	0.59 (0.501)	0.53 (0.779)
Age	40.25 (15.157)	42.79 (19.069)	43.80 (19.553)	44.13 (18.353)	40.75 (20.775)	42.94 (18.340)
Education	3.75 (1.292)	3.26 (1.327)	2.73 (1.223)	3.31 (1.580)	3.46 (1.503)	3.03 (1.383)
Income	6.43 (3.605)	4.79 (2.616)	3.71 (2.016)	4.50 (4.980)	5.73 (3.505)	4.60 (3.935)
Bid	68.07 (76.20)	197.53 (163.79)	96.17 (115.02)	145.94 (140.66)	139.45 (127.09)	122.72 (121.98)
Health Status[1]	2.12 (0.844)	2.26 (0.733)	2.73 (0.799)	1.88 (0.619)	2.18 (0.670)	2.44 (0.877)
Interview Interest	4.11 (0.925)	3.74 (0.991)	3.53 (0.516)	4.00 (0.679)	4.11 (0.801)	3.74 (0.657)
Interview Time	24.84 (7.258)	26.68 (8.420)	26.53 (10.225)	29.79 (7.934)	24.19 (9.835)	28.66 (9.110)

Table 3.0
Bayesian Estimates of the Endogenous Switching Tobit Model I

WTP Decision Model Variables	**With Imputed Values**		**With Unresolved Cases Excluded**	
	Coefficients	**Standard Deviation**	**Coefficients**	**Standard Deviation**
Intercept	0.046 ns	0.820	0.228ns	0.827
Cancer Experience	0.917**	0.134	1.046**	0.156
Health Status	1.692**	0.199	1.685**	0.203
Ln Age	0.301 ns	0.193	0.367 ns	0.197
Health and Ln Age	-0.495**	0.059	-0.487**	0.060
Ln Income	0.758**	0.181	0.789**	0.186
Ln Income^2	-0.072**	0.015	-0.080**	0.016
Ln Bid	-0.555**	0.037	-0.538**	0.036
True WTP Level Model Variables	**With Imputed Values**		**With Unresolved Cases Excluded**	
	Coefficients	**Standard Deviation**	**Coefficients**	**Standard Deviation**
Intercept	3.393**	0.852	3.569**	0.887
Education	0.287*	0.085	0.366**	0.104
Ln Number of Children	-0.043 ns	0.111	-0.101 ns	0.130
Ln Income	0.636**	0.200	0.707**	0.202
Ln Income^2	-0.074**	0.020	-0.087**	0.021
Ln Number of Children x Ln Income	0.005 ns	0.011	0.010ns	0.013
Ln Age	0.940**	0.215	0.957**	0.217
Ln Age^2	-0.276**	0.064	-0.274**	0.069

Parameter	Posterior Mean	Standard Deviation	Posterior Mean	Standard Deviation
Error Correlation	0.972**	0.005	0.975**	0.006
Sum of log CPO	-23.565		-22.698	

*indicates significant and ** highly significant at the 5% and 1% level respectively and [ns] indicates not significant.

Table 4.0
Mean WTP for Model I

Respondents	Posterior Mean	Standard Deviation	Posterior Mean	Standard Deviation
Demanders	143.818**	11.715	141.222**	13.021
Non-Demanders	27.680**	6.768	26.832**	7.333
Protesters (Yes)	143.506**	47.572		
Protesters (No)	21.518**	8.543		
Don’t Knows (Yes)	152.378**	33.161		
Don’t Knows (No)	23.541**	7.070		
Missing (Yes)	162.696**	15.161		
Missing (No)	19.153**	5.725		
True Zeros	-18.421**	8.95 e-13	-18.421**	8.95 e-13

Table 5.0
Classification of Unresolved Cases for Model I

Unresolved Cases	Mean Number Classified As	
	Yes	No
Protests	8.85	18.15
Don't Knows	12.63	28.36
Missing	1.11	1.89

Figure 3.0 Autocorrelation Functions of the MCMC Estimates

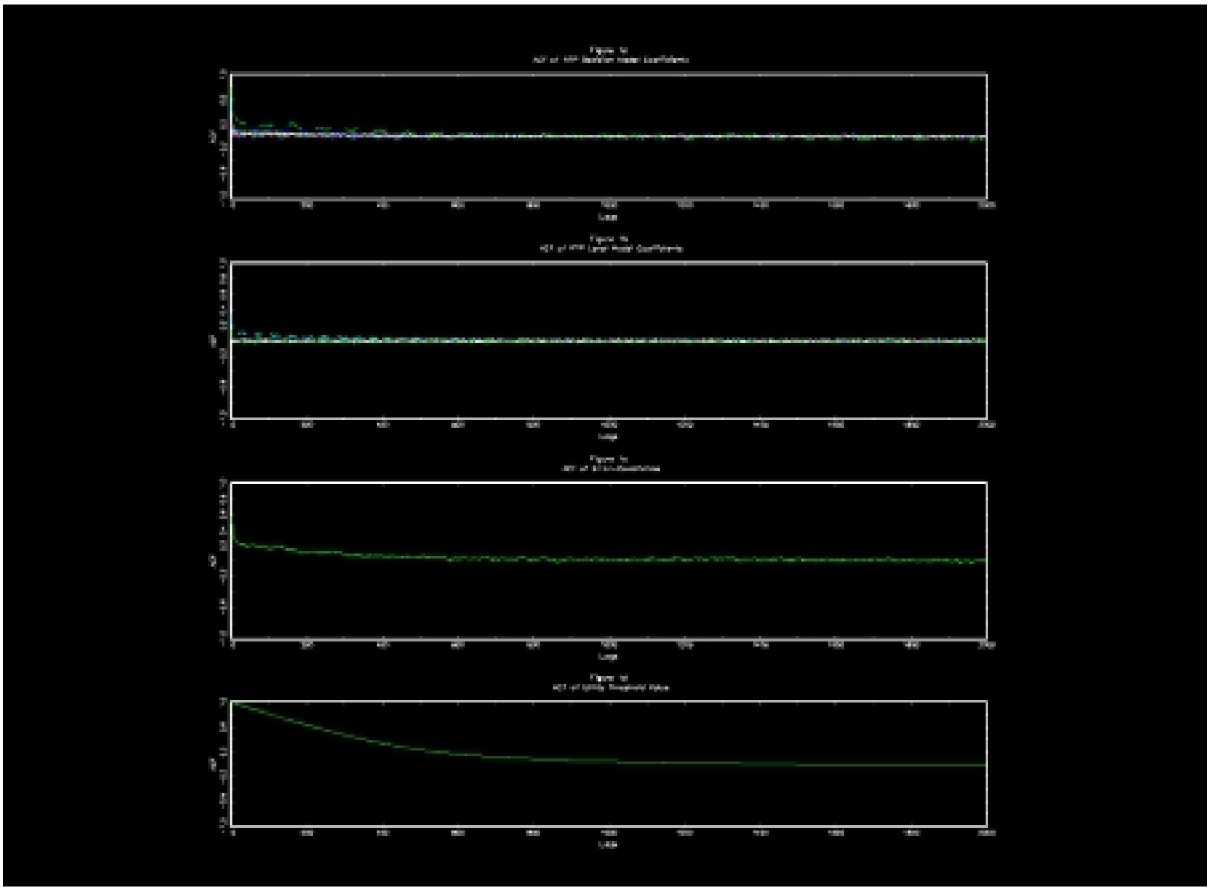

Table 6.0
Bayesian Estimates of the Endogenous Switching Tobit Model II

WTP Decision Model Variables	**With Imputed Values**		**With Unresolved Cases Excluded**	
	Coefficients	**Standard Deviation**	**Coefficients**	**Standard Deviation**
Intercept	-0.903 ns	0.821	-1.24 ns	0.839
Cancer Experience	1.200**	0.199	1.834**	0.207
Health Status	1.681**	0.200	1.696**	0.199
Ln Age	0.258 ns	0.175	0.248 ns	0.173
Health and Ln Age	-0.496**	0.058	-0.523**	0.061
Ln Income	0.570**	0.185	0.420*	0.185
Ln Income^2	-0.043**	0.015	-0.026 ns	0.015
Ln Bid	-0.239**	0.060	-0.104**	0.044
True WTP Level Model Variables	**With Imputed Values**		**With Unresolved Cases Excluded**	
	Coefficients	**Standard Deviation**	**Coefficients**	**Standard Deviation**
Intercept	3.810**	0.898	4.166**	0.931
Education	1.110**	0.314	2.174**	0.287
Ln Number of Children	-0.281 ns	0.174	-0.581**	0.196
Ln Income	0.751**	0.208	0.830**	0.212
Ln Income^2	-0.094**	0.023	-0.118**	0.025
Ln Number of Children x Ln Income	0.031 ns	0.018	0.064**	0.020
Ln Age	0.965**	0.216	0.970**	0.216
Ln Age^2	-0.430**	0.091	-0.639**	0.115

Parameter	**Posterior Mean**	**Standard Deviation**	**Posterior Mean**	**Standard Deviation**
Error Correlation	0.860**	0.062	0.654**	0.100
Utility Threshold Value	0.546**	0.052	0.725**	0.062
Sum of log CPO	-16.493		-16.875	

*indicates significant and ** highly significant at the 5% and 1% level respectively and ^{ns} indicates not significant.

Table 7.0
Mean WTP for Model II

Respondents	Posterior Mean	Standard Deviation	Posterior Mean	Standard Deviation
Demanders	152.530**	16.436	166.055**	20.658
Non-Demanders	31.893**	6.352	34.471**	6.266
Protesters (Yes)	206.717**	28.095		
Protesters (No)	14.225**	11.475		
Don't Knows (Yes)	214.522**	27.890		
Don't Knows (No)	35.231**	17.396		
Missing (Yes)	183.719**	51.603		
Missing (No)	5.068^{ns}	10.037		
True Zeros	1.00 e-8**	1.10 e-21	1.0 e-8**	1.099 e-21

Table 8.0
Classification of Unresolved Cases for Model II

Unresolved Cases	Mean Number Classified As	
	Yes	No
Protests	19.37	7.63
Don’t Knows	30.34	10.66
Missing	2.45	0.55

Appendix A

$$L(\theta / Y) = \prod_{i \in R_1} P[d^* > 0, w^* \geq b] \prod_{i \in R_{21}} P[d^* \leq 0, 0 < w^* < b] \prod_{i \in R_{22}} P[d^* \leq 0, w^* = 0]$$

$$= \prod_{i \in R_1} \{1 - \Phi(b - x_i' \beta)\} \quad \Phi\left[z_i' \gamma + \rho_{\eta\varepsilon} \left(\frac{w_i^* - x_i' \beta}{\sqrt{1 - \rho_{\eta\varepsilon}^{\ 2}}} \right) \right]$$

$$\prod_{i \in R_{21}} [\Phi(b - x_i' \beta) - 1 + \Phi(x_i' \beta)] \left\{ 1 - \Phi\left[z_i' \gamma + \rho_{\eta\varepsilon} \left(\frac{w_i^* - x_i' \beta}{\sqrt{1 - \rho_{\eta\varepsilon}^{\ 2}}} \right) \right] \right\}$$

$$\prod_{i \in R_{22}} \phi(x_i' \beta) \left[1 - \Phi\left[\frac{z_i' \gamma + \rho_{\eta\varepsilon}(w_i^* - x_i' \beta)}{\sqrt{1 - \rho_{\eta\varepsilon}^{\ 2}}} \right] \right]$$

Appendix B

MCMC Algorithm for EST Model 1

Imputation Step:

Generate the elements of the missing vector $\boldsymbol{M}_y^{(m)} = [\, d_i^M, d_i*, w*]$ from $[\boldsymbol{M}_y^{(m)} | \, ?_y^{(m-1)}]$ sequentially as follows:

1. For the unresolved cases or missing values generate the decision model dummy variable $d_i^{M\,(m)}$ from

   $[\, d_i^M \, |?^{(m?1)}] \sim Bernoulli\, [?\, (z_i'\, ?\,)]$.

2. For all cases generate the decision variable or net benefit score $d_i*^{(m)}$ from from a truncated normal

   $[\, d_i *|d_i ,?^{(m?1)}] \sim TN_{(??,0)}\, [z_i'\, ?\, ,1]$ with support $(-?\, , 0]$ if $d_i = 0$

   $[\, d_i *|d_i ,?^{(m?1)}] \sim TN_{(0,?\,)}\, [z_i'?\, ,1]$ with support $(0,?\,)$ if $d_i = 1$.

3. For the demanders, that is if $d_i = 1$, generate the true WTP w* in log-form from a truncated normal

$$[\, w_i *|d_i *^{(m)} ,d_i^{(m)} ,?^{(m?1)}] \sim TN_{[\ln b,U_L]}\, [x_i'\, ?^{(m?1)}\, ?\; ?_{??}^{(m?1)}(d_i *^{(m)}\, ?\, z_i'\, ?^{(m?1)}),(1?\; ?_{??}^{2(m?1)})]$$

   with support $[\ln b, U_L]$ *where* U_L is some reasonable upperbound like the $\ln(0.02I_{mp})$ or $\ln(m\, b)$ *where* I_{mp} is the midpoint of the income category the individual belongs to and m is an integer greater than or equal to 2 .

4. For the non-demanders, that is if $d_i = 0$, generate the true WTP w* in log-form from a truncated normal

$$[\, w_i * |\, d_i *^{(m)} ,d_i^{(m)} ,?^{(m?1)}] \sim TN_{(????,\ln b)}\, [\, x_i'\, ?^{(m?1)}\, ?\; ?_{??}^{(m?1)}(\, d_i *^{(m)}\, ?\, z_i'\, ?\,),(1?\; ?_{??}^{2(m?1)})]$$

   with support $(-?\;\; +?\; ,\ln b)$.

5. For the true zeros, that is if $d_i = 0$ and w*=0, generate the true WTP w* in

log-form from a truncated normal

$$[w_i * | d_i *^{(m)}, d_i^{(m)}, ?^{(m?1)}] \sim TN_{(? ?,? ? ??)}[x_i' ?^{(m?1)} ? ?_{??}^{(m?1)}(d_i *^{(m)} ? z_i' ?^{(m?1)}), (1 ? ?_{??}^{2(m?1)})]$$

with support in the neighborhood of -? , ? is some small positive number. Samples from the truncated normal are generated using Devroye's (1986) inversion method. Rejection method or Geweke's (1991) method may also be used.

<u>Posterior Step:</u>

6. Generate the variance-covariance matrix $?_?^{-1}$ from

$$[?^{?1} | ?^{(m?1)}, M_y^{(m)}] \sim W ??? ? n?, ??? R?^{?1} ? \sum_{i?1}^{n} ?^{(m?1)} ?^{(m?1)\,'} ??^{?1} ??$$

where $? ? ?Y_i ? X_i ?^{(m?1)}?$.

$?_? ? ?^{?1} ? ?^{?1}$, ? is a diagonal matrix with standard deviations of the random errors on the diagonal. This approach is simpler and more direct than Metropolis-Hastings which is more appropriate when the conditional distribution is intractable or not in closed form.

7. Generate ? from $[? | ?_?^{?1(m)}, M_y^{(m)}] \sim N(M_?, ?_?)$

which is a multivariate normal with mean

$$M_? ? ?_? ?V_?^{?1} A ? X' ?I_n ? ?_? ?^{?1} Y?$$

and variance-covariance

$$?_? ? ?V_?^{?1} ? X' ?I_n ? ?_? ?^{?1} X?^{?1}$$

Appendix C

MCMC Algorithm for EST Model 2

MCMC Algorithm for EST Model 2

<u>Imputation Step:</u>

Generate the elements of the missing vector

$\boldsymbol{M}_y^{(m)} = [\, d_i^M, (d_{th}*, d_i*)\, , w*]$ from $[\boldsymbol{M}_y^{(m)} | \, ?_y^{(m-1)}]$ sequentially:

1. For the unresolved cases or missing values generate

   the decision dummy variable $d_i^{\,M\,(m)}$ from

   $$[\, d_i^M \,|?^{\,(m?1)}] \sim Bernoulli[1 ? ? \; (d_{th} * ?z_i ' ? \;)/1 ? ? \; ??z_i ' ?\,]$$

   $d_i^M \; ? \; 2$ with probability p , $d_i^M \; ? \; 1$ with probability 1-p.

2. For all cases generate the threshold value $d_{th}*^{(m)}$ and the decision variable or

   net benefit score $d_i*^{(m)}$ from

   $$?\,?d_{th}*, d_i *?\,| d_i^{M\ (m)}, d_{th} *^{(m?1)}, ?^{(m?1)}?$$

   2a. $[\, d_{th} * |\, d_i^{M\,(m)}, d_{th} *^{(m?1)}, ?^{(m?1)}\,] \sim \; TN\, ?d_{th} *^{(m?1)}, ?_{d_{th}*}\,?$ with support $(0,?\,)$.

   $$d_{th}* \; = d_{th}*^{(m)} \qquad \text{if} \;\; U(\,0{,}1) \; ? \; min(1,? \;)$$

   $$= d_{th}*^{(m-1)} \qquad \text{otherwise}$$

   where

   $$? \;\; ? \; \frac{? \;\; ?d_{th} *^{(m?1)} \,/\, ?_{d_{th}*}\,?}{? \;\; ?d_{th} *^{(m)} \,/\, ?_{d_{th}*}\,?} \; ?_{d_i ?1} \frac{? \;\; ?d_{th} *^{(m)} \; ?z_i '?\,? ? \;\; ??z_i '?\,?}{? \;\; ?d_{th} *^{(m?1)} \; ?z_i '?\,? ? \;\; ??z_i '?\,?} \; ?_{d_i ?2} \frac{1? \; ? \;\; ?d_{th} *^{(m)} \; ?z_i '?\,?}{1? \; ? \;\; ?d_{th} *^{(m?1)} \; ?z_i '?\,?}$$

   2b. $[\, d_i * |\, d_{th} *^{(m)} d_i^{M(m)}, ?^{(m?1)}\,] \sim TN_{(??\,,0]}\,[\, z_i '\, ?^{(m?1)}, 1]$,if $d_i = 0$

   $[\, d_i * |\, d_{th} *^{(m)}, d_i^{M(m)}, ?^{(m?1)}\,] \sim TN_{(\,0\,, d_{th}*^{(m)}\,)}\,[\, z_i '\, ?^{(m?1)}, 1]$, if $d_i = 1$

   $[\, d_i * |\, d_{th} *^{(m)}, d_i^{M(m)}, ?^{(m?1)}\,] \sim TN_{[\,d_{th}*^{(m)}, ?\;)}\,[\, z_i '\, ?^{(m?1)}, 1]$, if $d_i = 2$.

   3a. For the demanders, that is if $d_i = 2$, generate the true WTP w* in log-form from the truncated normal

$$[\,w_i * \mid d_i *^{(m)}, d_{th} *^{(m)}, d_i^{M(m)}, ?^{(m?1)}\,] \sim$$

$$TN_{[lnb, U_L]}\,[\,x_i'\,?^{(m?1)}\;?\;?_{??}^{(m?1)}(\,d_i *^{(m)}\;?\;z_i'\,?^{(m?1)}\,), (1?\;?_{??}^{2(m?1)}\,)]$$

where $U_L = ln(0.02I_{mp})$ or $ln(m\,b)$

3b. For the non-demanders, that is if $d_i = 1$, generate the true WTP w* in log-form from the truncated normal

$$[\,w_i * \mid d_i *^{(m)}, d_{th} *^{(m)}, d_i^{M(m)}, ?^{(m?1)}\,] \sim$$

$$TN_{(????, lnb)}\,[\,x_i'\,?^{(m?1)}\;?\;?_{??}^{(m?1)}(\,d_i *^{(m)}\;?\;z_i'\,?^{(m?1)}\,), (1?\;?_{??}^{2(m?1)}\,)]$$

3c. For the true zeros, that is if $d_i = 0$ and w*=0, generate the true WTP w* in log-form from

$$[\,w_i * \mid d_i *^{(m)}, d_{th} *^{(m)}, d_i^{M(m)}, ?^{(m?1)}\,] \sim$$

$$TN_{(??, ????)}\,[\,x_i'\,?^{(m?1)}\;?\;?_{??}^{(m?1)}(\,d_i *^{(m)}\;?\;z_i'\,?^{(m?1)}\,), (1?\;?_{??}^{2(m?1)}\,)]$$

<u>Posterior Step:</u>

4.Generate the variance-covariance matrix $?_?^{-1}$ from

$$[\,?^{?1} \mid ?^{(m?1)}, M_y^{(m)}\,] \sim W\,??\;?\;n?, ???\,R?^{?1}\;?\;\sum_{i?1}^{n}\;?^{(m?1)}?^{(m?1)'}\,?^{?1}\,??$$

where $?\;?\;?Y_i\;?\;X_i?^{(m?1)}?$.

$?_?\;?\;?^{?1}??^{?1}$, $?$ is a diagonal matrix with standard deviations of the random errors on the diagonal.

5.Generate $?$ from multivariate normal

$$[\,?\;\mid ?_?^{?1(m)}, M_y^{(m)}\,] \sim N(M_?\,, ?_?\,)$$

with mean $M_?\;?\;?_?\;?V_?^{?1}A\;?\;X'?I_n\;?\;?_?\,?^{?1}\,Y?$

and variance-covariance $?_{?} \; ? \; ?V_{?}^{\,?1} \; ? \; X' ? I_n \; ? \; ?_{?} ?^{?1} X \; ?^{?1}$

References

Alvarez-Farizo, B., Hanley, N., Wright, R. and MacMillan, D. (1999) Estimating the benefits of agri-environmental policy: econometric issues in open-ended contingent valuation studies, *Journal of Environmental Planning and Management*, 42, 23-43.

Arrow, K. J., Solow, R., Leamer, E ., Portney, P., Radner, R., and Schuman, H . (1993) *Report of the NOAA Panel on contingent valuation*, *Federal Register*, 58, 4601-4614.

Bishop, R.C. and Heberlein, T. A . (1979) Measuring values of extra-market goods: are indirect measures biased?, *American Journal of Agricultural Economics*, 61,936-940.

Brox, J., Ramesh, K., and Strolley, K. (2003) Estimating willingness to pay for improved water quality in the presence of item nonresponse bias, *American Journal of Agricultural Economics*, 85, 414-428.

Calia, P. and Strazzera, E. (2001) A sample selection model for protests in contingent valuation analyses, *Statistica*, 61(3), 473-485.

Charlier, E., Melenberg, B. and van Soest, A. 2001) An analysis of housing expenditure using semiparametric models and panel data, *Journal of Econometrics*, 71:107.

Chib, S. and Hamilton, B. (2000) Bayesian analysis of cross-section and clustered data treatment models, *Journal of Econometrics* , 97, 25-50.

Copas, J. (1990) A likelihood approach to sample selectivity in multiple and binary regression models for comparative studies, *School of Mathematics and Statistics Working Paper 20/90 Univeristy of Birmingham.*

Cowles, M., Carlin, B. and Connett, J. (1996) Bayesian tobit modelling of longitudinal ordinal clinical trial compliance data with nonignorable misssingness, *Journal of the American Statistical Asscociation* , 91, 86-98.

Cowles, M. (1996) Accelerating Monte Carlo Markov Chain convergence for cumulative-link generalized linear models, *Statistical and Computing*, 6, 101-111.

Diggle, P. J., Liang, K. Y., and Zeger, S. L. (1994) *Analysis of Longitudinal Data*. Oxford University Press.

Donaldson, C., Jones, A., Mapp, T., and Olson, J. (1998) Limited dependent variables in willingness to pay studies: applications in health care, *Applied Economics*, 30, 667-677.

Greene, W. (1981) *Econometric Analysis*, 5th Edition, Prentice Hall, 2003.

Groothius, P. anad Whitehead, J. (2002) Does don't know mean no? Analysis of "don't know" responses in dichotomous choice contingent valuation questions *Applied Economics*, 34, 1935-1940.

Haener, M. and Adamowics, VW. (1998) Analysis of "don't know" responses to referendum contingent valuation questions, *Agricultural and Resource Economics Review* 29, 218-230.

Hanemann, W., 1984. Welfare evaluations in contingent valuation experiments with discrete responses: reply, *American Journal of Agricultural Economics*, 69, 332-341.

Heckman, J. (1976) The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models, *Annals of Economic and Social Measurement* 5: 475-92.

Heckman, J. (1979) Sample selection bias as a specification error, Econometrica 47, 153-61.

Horowitz, J. and Manski, C. (1998) Censoring of outcomes and regressors due to survey nonresponse: identification and estimation using weights and imputations, *Journal of Econometrics*, 84, 37-58.

Jorgensen, B.and Syme, G. (200) Protest responses and willingness to pay: attitude toward paying for stormwater pollution abatement, *Ecological Economics*, 33, 251-265.

Little, R., and Rubin, D. (1987) *Statistical Analysis with Missing Data*, Wiley, New York.

Little, R. (1988), Missing data adjustments in large surveys (with discussion), *Journal of Business and Economic Statistics*, 6, 287-297.
Little, R., and Wang Y. (1996) Pattern-mixture models for multivariate incomplete data with covariates, *Biometrics* 52, 98-111.
Little, R. and Rubin, D. (1999) Comment on adjusting non-ignorable dropout using semiparametric models by D.O. Scharfstein, A. Rotnisky and J. M. Robins, *Journal of the American Statistical Association* , 94, 1130-1132.
Little, R. and Hongyin, A. (2003) Robust likelihood based analysis of multivariate data with missing values, forthcoming *Statisca Sinisca*.
McClelland, E. and Whittington, D. (1994) The effect of information treatments on "don't know" responses in contingent valuation survey, USEPA, Oceans and Coast Protection Division.
Messonnier, M., Bergstrom, R., Cornwell, C., Teasley, R., and Cordell, H. (2000) Survey response-related biases in contingent valuation: concepts, remedies, and empirical application to valuing aquatic plant management *American Journal of Agricultural Economics*, 83, 438-450.
Nawata, K. and McAleer, M. (2001) Size characteristics of tests for sample selection bias:a Monte Carlo comparison and empirical example, *Econometric Reviews*, 20, 105-112.
Odejar, M. and Fahrmeir, L. (2002) Bayesian analysis of sample selection and endogenous switching regression models with random coefficients via MCMC methods, *Working Paper, SFB 386: Statistical Analysis of Discrete Structures, University of Munich*.
Rubin, D. (1976) Inference and missing data, *Biometrika* 63: 581-592.
Ryan, M ., Scott, D. and Donaldson, C. (2003) Valuing health care using willingness to pay:a comparison of the payment card and dichotomous choice methods, *Working Paper, University of Aberdeen, Health Economics Research Unit.*
Strazzera, E ., Scarpa, R ., Calia, P., Garrod, W . and Willis, K., (2003) Modelling zero values and protest responses in contingent valuation surveys, *Applied Economics*, 35, 133-138.
Wang, H. (1997) Treatment of "don't know" responses in contingent valuation surveys: a random valuation model, *Journal of Environmental Economics and Management*, 32, 219-232.