Journal of Agricultural and Resource Economics 26(2):491–507

Simulated Maximum Likelihood for Double-Bounded Referendum Models

Mary Riddel

Although joint estimation of referendum-type contingent value (CV) survey responses using maximum-likelihood models is preferred to single-equation estimation, it has been largely disregarded because estimation involves evaluating multivariate normal probabilities. New developments in the construction of probability simulators have addressed this problem, and simulated maximum likelihood (SML) for multiple-good models is now possible. This analysis applies SML for a three-good model under a double-bounded questioning format. Results indicate joint estimation substantially improves the variances of the parameters and willingness-to-pay estimates.

Key words: censored regression; contingent valuation; Geweke, Hajivassiliou, and Keane simulator; simulated maximum likelihood

Introduction

Contingent value (CV) surveys often seek to value multiple nonmarket goods within the same survey using the referendum approach. When multiple goods are valued, survey costs may be contained. Further, examining several goods simultaneously allows researchers to estimate substitution effects between potentially related goods. Past research has shown total willingness to pay (WTP) for related environmental goods is not simply the sum of the individual WTP for each good due to substitution effects between goods (Hoehn; Hoehn and Loomis; Carson; Carson and Mitchell). Thus, including several goods in a survey has clear advantages in terms of unbiased total value estimation and cost minimization.

Initially, researchers assumed the valuation responses were independent across individuals and goods, and estimated WTP for each good separately. Recent research has shown responses are often highly correlated between the goods, making the assumption of independence inefficient (Riddel and Loomis; Poe, Welsh, and Champ). When independence is ignored, parameter-variance estimates, and hence estimates of WTP variances, are overstated. The situation is analogous to a seemingly unrelated regressions model: when errors are correlated across goods, information is neglected that would allow more precise estimates of the covariance matrix.

Previous studies have reported that efficiency gains can be substantial from joint estimation. Riddel and Loomis show efficiency gains of 15% using a Monte Carlo simulation for a jointly estimated two-program, double-bounded model. Poe, Welsh and Champ suggest a similar result for joint estimation of two programs with a single-bounded model.

Mary Riddel is associate director of the Center for Business and Economic Research, and assistant professor of economics, University of Nevada, Las Vegas. The author thanks R. Keith Schwer, Thomas Carroll, John Loomis, Rennae Daneshvary, and an anonymous referee for helpful comments on this manuscript.

Review coordinated by Gary D. Thompson.

Minimum variance estimators allow policy makers to draw appropriate conclusions concerning the WTP for improvements in environmental quality. Using high-variance estimators could cause policy makers to erroneously infer that WTP for improved quality is not significantly greater than WTP for the existing level of environmental quality. Also, methodological tests for scope, substitution, and different survey designs require precise variance estimates for proper inferences. Estimates of the covariance between WTP for competing or substitute goods allow calculation of joint confidence regions around WTP values. These regions permit tests of hypotheses concerning preference order or substitutability of competing public goods.

The primary impediment to joint estimation has been the existence of multiple integrals in the likelihood equations. For the double-bounded referendum format with only one good, the log-likelihood function arising from the standard probit approach to estimation contains four bivariate normal probabilities. Expanding this model to only two goods means including quadravariate normal probabilities in the likelihood function.

Cameron's latent variable approach reduces the order of integration substantially, but increases the number of integrals to be estimated. For Cameron's approach, estimating a double-bounded, two-good model entails calculating bivariate normal integrals. These can be estimated readily with numerical methods. However, a three-good model involves evaluating trivariate normal integrals, and, because of the high order of integration, numerical methods for evaluating the integrals break down and thus preclude the use of standard maximum-likelihood techniques for these models.

New developments in Monte Carlo simulation have led to a new class of simulated maximum-likelihood models for solving many of the computational problems associated with jointly estimating censored distributions like those arising from dichotomous-choice referendum surveys. Simulated maximum likelihood (SML) models, based on work by Lerman and Manski, are quickly gaining ground in solving many problems which were previously intractable. The underlying idea behind SML is quite simple. One estimates the multivariate probabilities involving high-dimensional integrals using a probability simulator. For the popular simulator developed by Geweke, Hajivassiliou, and Keane (GHK), the multivariate probabilities are written as a sequence of conditional probabilities and simulated recursively, resulting in a simulated probability estimate. Given the simulated probabilities, standard optimization techniques are used to choose a parameter vector that maximizes the probability of obtaining the set of responses observed. In other words, conditioning on the simulated probabilities, one uses conventional maximum-likelihood techniques to estimate the model parameters.

The SML method is surprisingly simple to employ. Constructing programs to generate parameter estimates and the covariance matrix is relatively straightforward in a statistical programming language, such as GAUSS or STATA, containing a maximization routine. And, assuming correct model specification and relatively well-conditioned data, the models are quite well behaved.

The purpose of this study is to present the use of SML techniques when faced with intractable joint integrals regularly encountered in referendum-type survey data. The SML is based on the smooth recursive probability simulator developed in separate works by Geweke, Hajivassiliou, and Keane (GHK). Given the simulated probabilities, SML is used to estimate WTP for multiple goods under the double-bounded questioning format where multiple integrals typically arise.[1]

[1] See the special issue of *Review of Economics and Statistics*, volume 76, November 1994, devoted to probability simulators and their applications, for an in-depth discussion of different SML procedures.

Models for Contingent Valuation Survey Data

Consider the problem of estimating WTP values when several goods are included in a survey. We can specify a system of equations for the WTP amounts as a function of systematic and random components for individual i:

$$
\begin{aligned}
WTP_{i1} &= \mathbf{b}_1'\mathbf{X}_{i1} + e_{i1}, \\
WTP_{i2} &= \mathbf{b}_2'\mathbf{X}_{i2} + e_{i2}, \\
&\vdots \\
WTP_{iK} &= \mathbf{b}_K'\mathbf{X}_{iK} + e_{iK},
\end{aligned} \tag{1}
$$

where K equals the number of goods under consideration, WTP_{ik} is the willingness to pay for the kth good by the ith person, $\mathbf{b}_k'\mathbf{X}_{ik}$ is a linear combination of systematic components of WTP_{ik}, and e_{ik} is a random error associated with the ith individual and the kth good. The length of $\mathbf{b}_k$ may be different for each of the k goods, depending on what independent variables influence WTP for that good.

Error correlation may arise whenever there are linkages among the stochastic components of the WTP distributions. Mathematically, this reduces to $E[e_j e_k] \neq 0$ for $j \neq k$.[2] Conceptually, this may occur when hard-to-obtain, individual-specific characteristics, such as proxies for tastes and preferences, are not included in the empirical model. For instance, if two goods being valued in the survey are substitutes or complements, error correlation will be nonzero.

Another potential source of error correlation is unobserved attributes of the respondents that lead to nonzero errors across the equations. For example, the preference function leading an individual to prefer one type of good may cause the individual to value another good less. Of course, surveys of reasonable length can only elicit so much information about respondents; thus error correlation, and the resulting inefficiencies, are almost certainly troublesome.

For contingent valuation surveys employing the referendum approach, WTP is a latent variable: only bounds on the actual WTP amounts are observed.[3] Recognizing this, Cameron suggests an estimation approach based on censored distributions.[4] Under her model, the bid amounts, together with the responses, define a range of WTP. For example, individual i may be asked if she would be willing to pay t_i for some nonmarket good. If she responds yes, t_i defines the lower bound for her WTP. If she responds no, then t_i defines the upper bound of her WTP. A set of individual or good-specific regressors,

---

[2] If we assume normal error terms with contemporaneous correlation between the K error terms, so that $E[e_j e_k] \neq 0$ for $j \neq k$, the underlying model is a seemingly unrelated regression (SUR) model (Greene).

[3] In the single-bounded referendum model, respondents are offered a bid amount and asked if they would be willing to pay that amount. They respond yes or no, and the bid amount becomes a lower or upper bound, respectively, to the latent WTP value. In the double-bounded model, a second bid is proposed depending on the response to the first bid. A yes response to the first question is followed by a stepped-up bid amount, and a no response is followed by a lower bid amount. The responses to the two questions offer up to two bounds for each latent WTP.

[4] An alternative estimation method to that presented here is the probit approach following Hanemann and Kanninen. The probit approach can be expanded to the multi-good setting, but at great cost in terms of computation time and precision. For a two-good model with a follow-up question, the likelihood function involves the quadrivariate normal because each of the two questions for the four goods must be included as a random variable. In general, in the probit approach for K goods with a double-bounded questioning format, the order of integration necessary to estimate the model parameters is $2K$. Cameron's approach, on the other hand, involves only K-order integration when K goods are studied. Although the GHK estimator will handle either approach, the computing time may be excessive for higher-order integrals. For that reason, this article focuses on Cameron's latent variable approach due to its superior computational properties.

$\mathbf{X}_i$, together with the parameters, $\mathbf{b}$, define the conditional mean WTP. For the one-good, single-bid model, assuming a normal error distribution, the following likelihood function may be estimated using standard optimization techniques where $y_i = 1$ if a yes response is given and zero otherwise:

$$(2) \qquad \text{Log}\,\text{L} = \sum_{i=1}^{n} \{y_i \log[\Phi((t_i - \mathbf{b}'\mathbf{X}_i)/\sigma)] + (1 - y_i)\log[1 - \Phi((t_i - \mathbf{b}'\mathbf{X}_i)/\sigma)]\}.$$

Cameron's approach easily generalizes to the dichotomous choice with follow-up format. Here, there are two elicitation questions, with four outcomes: (YES, YES), (YES, NO), (NO, YES), and (NO, NO). The probabilities of these four outcomes may be defined as p^{11}, p^{10}, p^{01}, and p^{00}, respectively, where p^{11} is the probability the individual responds yes to both questions, and p^{10} is the probability of a yes to the first question and no to the second, and so on.

Accordingly, the probabilities of observing each outcome are:

$$(3) \qquad \begin{aligned} p^{11} &= \text{pr}(\text{bid } u \le WTP), \\ p^{10} &= \text{pr}(\text{bid } r \le WTP < \text{bid } u), \\ p^{01} &= \text{pr}(\text{bid } d \le WTP < \text{bid } r), \\ p^{00} &= \text{pr}(WTP \le \text{bid } d), \end{aligned}$$

where bid r is the initial bid, and bid u and bid d are the step-up and step-down bid offers, respectively. Assuming normally distributed errors with the normal cumulative distribution function defined as $\Phi(\cdot)$, the contribution to the log-likelihood function from the ith individual is specified as follows (Alberini; Cameron):[5]

$$(4) \qquad \begin{aligned} \text{Log}\,\text{L}_i = {} & I_i^{11} * \log(\Phi(z_{2i})) + I_i^{10} * \log(\Phi(z_{2i}) - \Phi(z_{1i})) \\ & + I_i^{01} * \log(\Phi(z_{1i}) - \Phi(z_{2i})) + I_i^{00} * \log(1 - \Phi(z_{2i})), \end{aligned}$$

where $z_{1i} = (c_{1i} - \mathbf{b}_1'\mathbf{X}_{1i})/\sigma$, c_{1i} is the first bid offer for the ith respondent, $z_{2i} = (c_{2i} - \mathbf{b}_1'\mathbf{X}_{1i})/\sigma$, c_{2i} is the follow-up bid offer for the ith respondent, and I_i^{mn} denotes an indicator variable corresponding to the category of the observed response. The estimate of individual i's willingness to pay for the good is calculated as $\mathbf{b}_1'\mathbf{X}_{1i}$.

Cameron's approach also easily generalizes, in theory, to estimating WTP for multiple goods included in a survey. For two goods, we have two sets of probabilities: $p_1 = \{p_1^{11}, p_1^{10}, p_1^{01}, p_1^{00}\}$ and $p_2 = \{p_2^{11}, p_2^{10}, p_2^{01}, p_2^{00}\}$, where the subscript $j = 1, 2$ denotes good 1 and good 2. In this case, there are 16 possible combinations of two responses for each of the two goods. To allow for error correlation, it is necessary to use a bivariate cumulative distribution function (CDF). The likelihood function contains terms from the joint CDF for each possible outcome. The probability of the respondent answering yes to the initial and follow-up bids for both goods is written as:[6]

$$(5) \qquad \begin{aligned} p_{1i}^{11} \cap p_{2i}^{11} &= \Pr(z_{21i} < WTP_{1i} < \infty;\ z_{22i} < WTP_{2i} < \infty) \\ &= \Phi(z_{21i}, \infty;\ z_{22i}, \infty;\ \rho), \end{aligned}$$

[5] This model, referred to as the interval-data model, assumes the WTP distribution for good k does not change between the initial and follow-up questions. Alberini shows that as long as the correlation between the underlying WTP distributions is high (>0.6), then the interval-data model is more efficient than the bivariate model proposed by Cameron.

[6] The likelihood function for the two-good, double-bounded model is reported in appendix A.

where $z_{21i} = (c_{21i} - \mathbf{b}_1'\mathbf{X}_{1i})/\sigma_1$, c_{21i} is the follow-up bid offered to the ith respondent for the first good, $z_{22i} = (c_{22i} - \mathbf{b}_2'\mathbf{X}_{2i})/\sigma_2$, c_{22i} is the follow-up bid offered to the ith respondent for the second good, and σ_1, σ_2, and ρ are the standard deviations and the correlation for the errors in the first and second equations of equation system (1) (e_1 and e_2, respectively).

Expanding the model to three goods necessitates calculating N trivariate normal probabilities for each iteration of the maximization process, where N is the sample size. With three goods, there are $4^3 = 64$ possible trivariate integrals. For instance, if the response set is (YES, YES), (YES, NO), and (NO, YES) for the first, second, and third goods, respectively, the joint probability is denoted by:

$$(6)\qquad p_{1i}^{11} \cap p_{2i}^{10} \cap p_{3i}^{01} = \Pr(z_{21i} < WTP_{1i} < \infty;\ z_{12i} < WTP_{2i} < z_{22i};\ z_{23i} < WTP_{3i} < z_{13i}) \\ = \Phi(z_{21i}, \infty;\ z_{12i}, z_{22i};\ z_{23i}, z_{13i};\ \rho_{12}, \rho_{13}, \rho_{23}),$$

where ρ_{ij} is the error correlation between e_i and e_j.

Standard maximum-likelihood techniques cannot handle terms containing high orders of integration for the normal or lognormal because the multivariate integrals are either time consuming or impossible to calculate numerically. The issue may be side-stepped by using a distribution from the generalized extreme value (GEV) family of distributions such as the multinomial logit or the nested multinomial logit. The disadvantage of GEV models is that they impose a rigid correlation structure on the covariance matrix. For example, in the multinomal logit model with J alternatives, the disturbance terms, e_{ij}, are independent, identically distributed for all $j = 1, \ldots, J$.

If we abandon GEV models to allow for a flexible correlation structure, it is necessary to estimate multiple integrals of the chosen error distribution. These integrals are invariably troublesome. The logistic distribution, though easier to integrate, is not appropriate for modeling joint WTP because it constrains the correlation coefficients to unity (Cameron and Quiggin; Hanemann and Kanninen). The normal distribution is preferred to the logistic because it allows for positive correlation between WTP amounts.

Nevertheless, a closed form does not exist for the normal cumulative distribution function, so either numerical or simulation methods are necessary to calculate probabilities from the joint CDF. Numerical methods work quite well for calculating bivariate normal integrals, but become frustratingly slow for trivariate integrals (Greene, p. 183). For higher orders of integration, numerical methods are not computationally efficient (Greene, p. 183).

Because of these problems, standard maximum-likelihood techniques are not appropriate for modeling censored joint distributions with more than three variables, and are cumbersome for censored WTP models with three variables. SML techniques, although still computationally time consuming, easily generalize to the censored multiple-good framework. For this reason, we turn to a discussion of SML techniques for joint estimation of WTP distributions.

The GHK Simulator and Simulated Maximum Likelihood

The difficulties associated with calculating high-order integrals have spurred a flurry of research in Monte Carlo integration and simulation techniques. Monte Carlo integration techniques involve repeated sampling from a distribution related to the one in question. The sample average of the simulated probabilities converges to the desired probability of the distribution. In the most general method proposed by Metropolis and Ulam, we are interested in evaluating the integral for some general function $f(y)$ continuous in the range a to b:

$$F(y) = \int_a^b f(y) = \int_a^b v(y)g(y)\,dy. \tag{7}$$

If we assume $K = \int_a^b g(y)\,dy$, then dividing by K will create the probability function $h(y) = g(y)/K$ with related cumulative density function $H(y)$. Here, $H(a) = 0$ and $H(b) = 1$, so that the limits of the new probability function H are exactly the limits of the integral $F(y)$. Substituting into (7),

$$F(y) = K\int_a^b v(y)\frac{g(y)}{K}\,dy = KE_{h(x)}[f(y)].$$

Because $h(y) = g(y)/K$ is a probability density function, $F(y)$ is just K times the expected value of function $v(y)$ with y having distribution function $H(y)$.

Drawing $i = 1$ to R random samples individually designated as y_{ih} from the density function $h(y)$, then by the Strong Law of Large Numbers:[7]

$$\hat{F}(y) = \frac{K}{R}\sum_{i=1}^{R} f(y_{ih}) \xrightarrow{\text{m.s.}} F(y), \tag{8}$$

where m.s. represents convergence in mean square as the sample size, R, becomes large.

The challenge is to find some decomposition of the probability density function (pdf), $f(y) = v(y)g(y)$, so that $g(y)/K$ is a known pdf from which one can sample in a straightforward way. A simple application of this approach could, in theory, be used to calculate trivariate normal integrals such as $\Phi(a_1 < y_1 < b_1;\ a_2 < y_2 < b_2;\ a_3 < y_3 < b_3)$, where Φ is the normal CDF. One may generate R random draws from the trivariate normal, and score a 1 when the values fall in the specified intervals and a zero otherwise. The sum of the scores divided by R is the estimated probability.

While this simple approach is intuitively appealing, sample sizes allowing for good approximations are large and computing time is excessive (Greene, pp. 179–85). As a result, other less computer-intensive approaches have been developed for multivariate normal distributions. Probably the most successful, in terms of computer time necessary and convergence properties of the estimators, is the GHK smooth recursive simulator (Geweke; Hajivassiliou; Keane).

According to Greene, the K-variate normal integral with covariance matrix Σ can be estimated as an average over repeated draws from truncated joint normal distributions so that:

$$\Phi(a_1 < y_1 < b_1;\ a_2 < y_2 < b_2;\ \ldots;\ a_K < y_K < b_K) \approx \frac{1}{R}\sum_{r=1}^{R}\prod_{k=1}^{K} Q_{rk},$$

where the Q_{rk} are univariate truncated normal probabilities with a K-variate joint distribution $\Pi_{k=1}^{K} Q_{rk}$. These Q_{rk} can be calculated using the following k steps:[8]

---

[7] See Robert and Casella (p. 75) for a discussion of the properties of $h(y)$ which are necessary to ensure this result.

[8] This explanation follows the discussion by Greene (pp. 181–85).

- STEP 1. Calculate $Q_{r1} = \Phi(b_1/m_{11}) - \Phi(a_1/m_{11})$, where **M** is the lower triangle matrix obtained from the Cholesky decomposition of Σ with elements m_{kl}.

- STEP 2. Generate w_{r1} from the standard normal distribution truncated at a_1/m_{11} on the left and b_1/m_{11} on the right.[9] Using these, calculate:

$$A_{r2} = (a_2 - m_{11}w_{r1})/m_{22},$$
$$B_{r2} = (b_2 - m_{11}w_{r1})/m_{22},$$
$$Q_{r2} = \Phi(B_{r2}) - \Phi(A_{r2}).$$

- STEP k. Generate w_{rk-1} from the standard normal distribution truncated at A_{rk-1} on the left and B_{rk-1} on the right, and then calculate:

$$A_{rk} = \left(a_k - \sum_{l=1}^{k-1} m_{kl}w_{rl}\right)/m_{kk},$$
$$B_{rk} = \left(b_k - \sum_{l=1}^{k-1} m_{kl}w_{rl}\right)/m_{kk},$$
$$Q_{rk} = \Phi(B_{rk}) - \Phi(A_{rk}).$$

Steps 1 through k are performed R times, one for each of the R sampling points. The final estimate of the probability is calculated as:

$$\hat{\Phi}(a_1 < x_1 < b_1;\, a_2 < y_2 < b_2;\, \ldots;\, a_K < y_K < b_K) = \frac{1}{R}\sum_{r=1}^{R}\prod_{k=1}^{K} Q_{rk}. \tag{9}$$

The GHK simulator has superior statistical properties to those of other probability simulators proposed by Stern; Lerman and Manski; and Breslaw, and those based on the Gibbs sampling algorithm. Discontinuous frequency simulators, like those proposed by Lerman and Manski, provide consistent estimates only if the number of draws goes to infinity, an unfortunate property for a probability simulator. Like Stern's simulator, the GHK simulator is unbiased and consistent (Borsch-Supan and Hajivassiliou). And perhaps more importantly, the smoothness of the GHK and Stern's simulator keeps the number of draws required for good approximation in the range of 20 to 40, well below the 10,000+ draws or so required by other simulators (Geweke, Keane, and Runkle).

One drawback of Stern's simulator is that when error correlation becomes high, the variance of the simulated probabilities increases, a problem not shared with the GHK simulator. According to Borsch-Supan and Hajivassiliou, the GHK simulator has substantially smaller variance than Stern's method when errors are highly correlated. Because past research has shown high correlation between bid responses over goods, the GHK simulator will have superior properties to Stern's simulator for evaluating WTP from referendum data (Riddel and Loomis; Cameron and Quiggin).

Breslaw has suggested a simulator which uses the GHK simulator to estimate a line integral between a desired point on the CDF and an initial point close to the desired point. This simulator adds precision to the estimates, but also computational time. Breslaw recommends this simulator for situations when the inaccuracy of the simulator

[9] Hajivassiliou has suggested the following method for generating a draw from the normal constrained to lie between a and b. Let $U \sim \text{unif}(0, 1)$. Then $W = \Phi^{-1}[(\Phi(b) - \Phi(a)) * U + \Phi(a)] \sim N(0, 1)$.

causes convergence problems. Convergence problems were not experienced in the following application, so the standard GHK method was used.

The Gibbs sampler has also been proposed for evaluating multivariate normal integrals. A simulation experiment by Geweke, Keane, and Runkle found the Gibbs sampling approach had an advantage for models with large numbers of parameters (>20). The GHK simulator is preferred when the model contains fewer parameters. The following model contains 19 parameters; thus we chose the GHK simulator for this reason and due to the comparative ease of coding the GHK simulator relative to the Gibbs sampling algorithm.

Simulated Maximum Likelihood

Estimating the parameters of a three-good, double-bounded model is an imposing task because it involves evaluating as many trivariate normal CDFs as there are observations for every iteration of the maximization process. SML has the potential to simplify these calculations substantially. For the four-good model, SML is currently the only practical option.[10] To create the simulated likelihood function, we replace $\Phi(\cdot)$ with $\hat{\Phi}(\cdot)$ in the likelihood function. For example, for the one-good, double-bounded model, individual i's contribution to the SML equation is specified as:

$$(10) \qquad \text{Log}\,\text{L}_i = I_i^{11} * \log(\hat{\Phi}(z_{2i})) + I_i^{10} * \log(\hat{\Phi}(z_{2i}) - \hat{\Phi}(z_{1i})) + I_i^{01} * \log(\hat{\Phi}(z_{1i}) - \hat{\Phi}(z_{2i})) + I_i^{00} * \log(1 - \hat{\Phi}(z_{2i})).$$

Appendix A includes the SML function for the two-good, double-bounded model. A GAUSS program for the SML of the three-good, double-bounded model is available from the author upon request.

The GHK simulator provides unbiased estimates of the probability being estimated. However, the simulated log-likelihood function derived from the GHK simulator is not unbiased because of the nonlinearity inherent in the logarithm. In fact, the log-likelihood, and the estimated model parameters, may be biased for any positive variance of the probability simulator. The parameter bias increases at an approximately linear rate with the variance of the simulator. For this reason, a low-variance simulator, such as the GHK, is essential for SML parameter estimation. Nevertheless, Borsch-Supan and Hajivassiliou claim the bias of the simulated log-likelihood function, and hence the bias in the model parameters, should be small in a well-specified model. Using simulations, they show that for $R = 20$ the bias in the log-likelihood function is minor.

Confidence Regions for Expected WTP

Developing confidence regions around estimated WTP amounts is of special interest in the multiple-good survey. Often, the point of offering multiple goods within a survey is to decide whether one good is preferred to another. Alternatively, one may be interested in testing for substitution effects among goods in the survey. Regardless, when WTP

[10] The trivariate normal CDF can be evaluated using numerical techniques, although the task is computationally time consuming. The GHK, even using small numbers of draws, is also time consuming, but was found to be faster than numerical procedures for this application.

is estimated for each good individually, confidence regions surrounding mean WTP estimates are inaccurate because the correlation structure of the entire coefficient matrix is not considered.[11] When correlation among the model parameters across equations is present, the confidence region arising from joint estimation is necessarily smaller than one produced from single-equation estimation. Hence, joint inference using individually estimated mean WTP values may provide erroneous results. Derivation of the confidence region surrounding jointly estimated mean WTP for two goods is presented in appendix B.

The smaller confidence region gained through joint estimation becomes most important in the context of estimating differences in the expected WTP. If the ellipse contains the line where $WTP_1 = WTP_2$, then we can infer that the values of the two goods are not significantly different. Because the area of a correlated confidence ellipsoid is smaller than the rectangular region inferred when correlation is ignored, it is less likely to contain points on the line where $WTP_1 = WTP_2$. As such, the tighter confidence regions provide more discernment for evaluating WTP differences among goods.

An Application

To clarify the method and provide concrete examples of applications of the joint estimation method using SML, WTP was estimated for three programs designed to reduce fire hazard in California and Oregon's spotted owl habitats located in old growth forests. The data used were obtained from a survey conducted in 1995. A professional telephone survey company was used and repeated efforts were made to complete contacts. Respondents were initially contacted and asked to participate. If they agreed, they were sent an information booklet and an interview was scheduled for a convenient time. During the interview, respondents were questioned about demographic characteristics and double-bounded referendum questions were used to obtain bounds on WTP for each respondent.

Surveyors used random-digit dialing to contact 709 California households. Of these, 499 households agreed to an in-depth interview and were mailed booklets containing information about the proposed programs. The respondents were then telephoned again and asked double-bounded WTP questions for these programs. Demographic questions were also posed at this time. A total of 358 interviews were completed. From those completed interviews, 343 were used for the analysis; the remaining responses were dropped due to missing data in the question responses. This provides a final response rate of 48%.

Three programs were evaluated in the survey. The first program was designed to reduce high-intensity fires in California spotted owl habitats through fire-hazard reduction, early fire detection, increased fire protection, and larger fire control response than the status quo. The second program was targeted at Oregon spotted owl habitat, using the same fire-control methods. The third program (termed the combined program) would fund fire-hazard reduction in both California and Oregon.

Information booklets were mailed to the respondents prior to their telephone interview.[12] The booklets contained textual and visual information concerning the different

[11] The coefficient matrix is composed of the coefficient vectors for each of the K goods.

[12] Copies of the interviewer script and program booklet are available from the author on request.

programs. Respondents were asked to refer to these booklets to assist them in formulating their WTP for the different programs.[13]

WTP was elicited through the standard double-bounded approach. For the California program, respondents were asked the following:

> Q 24. Thinking about program B, which reduces the proportion of high-intensity fires like program A, and also includes a 20 percent reduction in the acreage of old growth forest that burns every year: If program B were the only program available and your household was asked to pay $[*fill bid*] each year to help pay for program B, would you pay this amount?
>
> <1> Yes [go to Q 24a]
> <2> Refused [go to Q 24b]
>
> Q 24a. What if the costs of program B were higher: Would your household pay $[*fill bid*] to have the program?
>
> Q 24b. What if the costs of program B were lower: Would your household pay $[*fill bid*] to have the program?

Estimation and Results

To illustrate the SML, the mean WTP of each of the three programs was first estimated individually using Cameron's maximum-likelihood approach. Next, the SML technique was used to jointly estimate the WTP values. The lognormal distribution is used because, unlike the normal distribution, it does not imply a range of negative WTP values. Rather, WTP values arising from lognormal distribution will be strictly positive. The bid values are transformed to the lognormal distribution by taking the natural log of the bid values and estimating the equation in appendix A using logged bid values (see Hanemann and Kanninen). Thus the system of equations is:[14]

$$\ln(WTP_{ik}) = \beta_k' \mathbf{X}_{ik} + e_{ik} \quad \text{for } k = 1, 2, \ldots, K. \tag{11}$$

Joint estimation results, together with the results from the single-equation estimation, are reported in table 1. In the single-equation models, two variables were significant at the $\alpha = 0.05$ level in explaining variation in WTP for the California and combined programs. These are *OGEXIST*, an integer variable measuring the importance of the existence of old growth forests to the respondent, and *EQIMP*, another integer variable measuring the importance of environmental quality. These variables were also significant in the Oregon model, but in addition, *DONS*, a dummy variable indicating whether the respondent had made any donations or contributions for wildlife or environmental protection in the last 12 months, is marginally significant at the $\alpha = 0.15$ level.

---

[13] The bid design was developed using responses from a single-bound dichotomous-choice CVM survey on northern spotted owls. Having multiple scenarios will affect the bid design if the follow-on scenario involves higher levels of environmental quality or a broader basket of environmental goods. In this case, there is a delicate balancing of increasing the bid for the second scenario, but not by more than the likely increment in value provided by the second scenario. If the bids on the second, better quality scenario are bumped up too much, the result could be more "no" responses on the better quality program.

[14] The Greek symbol, β, is used here to distinguish the coefficient matrix obtained by using the lognormal distribution from that estimated using the normal distribution for WTP.

Table 1. Estimated WTP Values, Standard Errors, and *p*-Values: Joint and Individual Estimation for Oregon, California, and Combined Programs

	Joint Estimation			Single Estimation		
Description[a]	Parameter	Std. Error	*p*-Value	Parameter	Std. Error	*p*-Value
Oregon:						
Intercept	-0.2803	0.6737	0.3387	-0.7024	0.8642	0.2103
OGEXIST	0.4887	0.1459	0.0004	0.4813	0.1923	0.0024
EQIMP	0.5387	0.1678	0.0007	0.6062	0.2743	0.0043
DONS	-0.0795	0.0500	0.1256	0.2058	0.2122	0.1545
σ_1	1.4141	0.0778	0.0000	1.4737	0.1029	0.0000
E(WTP)	***3.5763***	***0.0838***	***0.0000***	***3.4854***	***0.0962***	***0.0000***
California:						
Intercept	-0.2942	0.8793	0.3690	-0.7250	0.8286	0.1908
OGEXIST	0.5415	0.1580	0.0003	0.5452	0.1585	0.0003
EQIMP	0.4895	0.1835	0.0002	0.6877	0.2166	0.0008
σ_2	1.3144	0.0702	0.0000	1.4349	0.1001	0.0000
E(WTP)	***3.9066***	***0.0756***	***0.0000***	***3.7769***	***0.0927***	***0.0000***
Combined Program:						
Intercept	0.4835	0.5467	0.1882	0.1899	0.7020	0.3944
OGEXIST	0.5370	0.1183	0.0000	0.5409	0.1471	0.0002
EQIMP	0.4280	0.1561	0.0031	0.4790	0.2137	0.0800
σ_3	1.3363	0.0691	0.0000	1.3570	0.0836	0.0000
E(WTP)	***4.0877***	***0.0729***	***0.0000***	***3.9784***	***0.0862***	***0.0000***
Correlation Coefficients for Joint Estimation:						
ρ_{12}	0.7907	0.0233	0.0000	—	—	—
ρ_{13}	0.8862	0.0135	0.0000	—	—	—
ρ_{23}	0.8129	0.0185	0.0000	—	—	—

n = 343
Log-likelihood Oregon individual = -413.678
Log-likelihood California individual = -414.2934
Log-likelihood Combined Program = -657.89

[a] *OGEXIST* = an integer variable measuring the importance of the existence of old growth forests; *EQIMP* = an integer variable measuring the importance of environmental quality; *DONS* = a dummy variable indicating whether respondent had made any donations or contributions for wildlife or environmental protection in the last 12 months; σ_1, σ_2, and σ_3 = respective standard deviations for Oregon, California, and combined programs.

Given the coefficient matrix for the logged equations, β, one can estimate $E[\ln(WTP_k)] = \hat{\beta}_k' \bar{\mathbf{X}}_k$ for equation k. Single-equation estimates of $E[\ln(WTP_k)]$ for the California, Oregon, and combined programs are 3.7769, 3.4854, and 3.9784, respectively.[15] Using the relationship between the lognormal and the normal, $E[WTP]$ = \$122.29, \$96.67, and \$134.17 for the California, Oregon, and combined programs, respectively. California residents are more likely to express high WTP amounts for California programs because they are the direct beneficiaries in terms of use-value of the improvements. The WTP amount for the combined program is substantially less than the sum of the WTP amounts for Oregon and California programs provided separately, suggesting the possibility of

[15] Using the formula $\mu = \exp[\mu_{\ln} + \frac{1}{2}\sigma_{\ln}^2]$, where $WTP \sim N(\mu, \sigma^2)$ and $\ln(WTP) \sim \text{lognormal}(\mu_{\ln}, \sigma_{\ln}^2)$, one is able to analytically transform the mean of a lognormal distribution to that of a normal distribution.

substitution effects between the Oregon and California programs. Single-equation estimates do not provide the covariance structure to properly test for substitution between the California and Oregon programs, so we turn to the jointly estimated WTP model for such a test.

The first numeric column in table 1 gives the parameter estimates and estimated $E[\ln(WTP)]$ for the three programs estimated jointly using SML. The correlation of the responses to the programs is quite high: 0.88 between the Oregon and combined programs and 0.81 between the California and combined programs, and 0.79 between the Oregon and California programs. The same set of independent variables is significant in the California and combined programs. However, in the Oregon program, the dummy variable *DONS* (representing respondent donations to environmental causes within the past 12 months) is also significant.

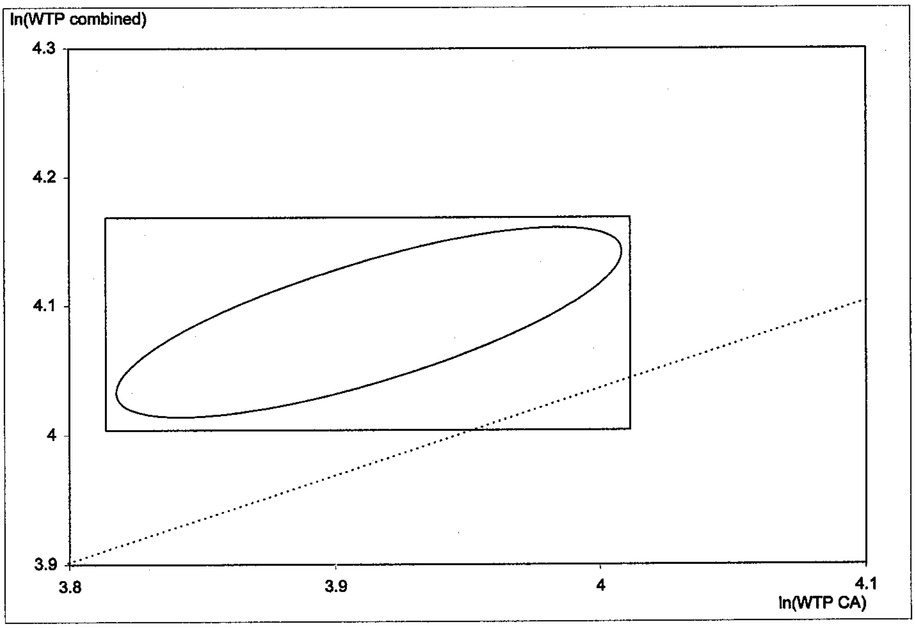
In an uncensored SUR model, the different regressor sets potentially give rise to smaller standard errors for the parameters of the jointly estimated equation. An analogous outcome is observed in the censored SUR model in table 1:[16] standard errors of the coefficients are 15 to 22% less under joint estimation than under single-equation estimation, underscoring the precision gains that may be achieved through joint estimation. Similar gains are seen in the precision of the $E[\ln(WTP)]$ estimates.

Under joint estimation, the WTP estimates do not change by much; confidence intervals around all the parameters under joint estimation contain the single-equation point estimates. The $E[\ln(WTP)]$ amounts for the California, Oregon, and combined programs are 3.5763, 3.9066, and 4.0877, respectively. Translated into dollar amounts, the California, Oregon, and combined programs are worth \$117.97, \$97.14, and \$145.55, respectively. Thus, the mean WTP amounts estimated using the SML are not appreciably different from those estimated using single-equation estimation; the primary advantages of SML in this application are smaller standard errors and the facility of estimating the joint confidence regions.

The confidence region for $E[\ln(WTP \text{ California})]$ and $E[\ln(WTP \text{ combined})]$ is plotted in figure 1. The 45° line does not pass through the confidence region, implying the hypothesis that the expected WTP values for the two programs are equal can be rejected at the 5% level. The rectangle represents the confidence region associated with single-equation estimation. As expected, joint estimation provides smaller confidence regions for testing the hypothesis that the two programs have equal $E(WTP)$ values.

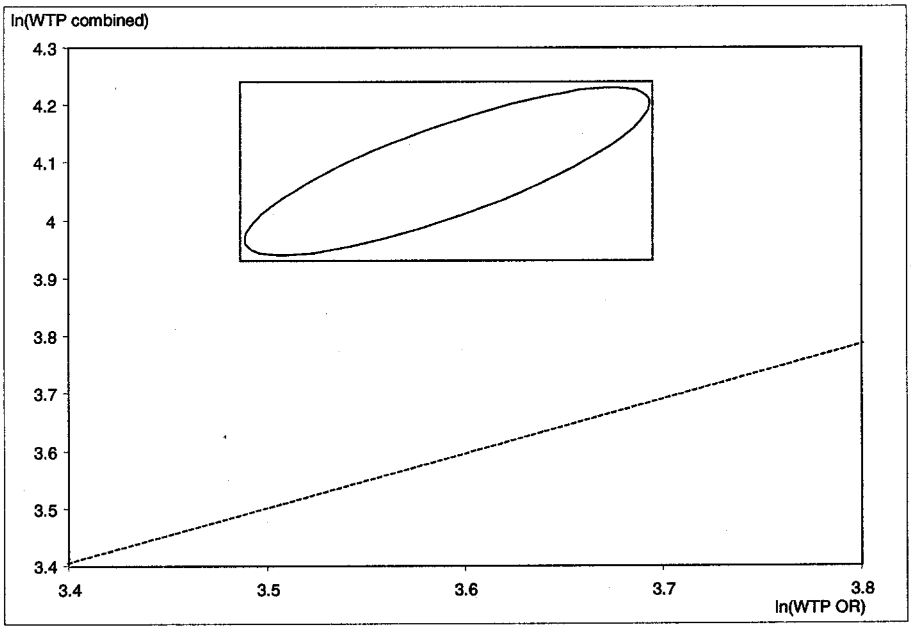
The 95% confidence region for the Oregon and combined programs is illustrated in figure 2. Because the confidence ellipse under joint estimation does not contain the 45° line, we can reject the hypothesis that the WTP values for the two programs are equal. The region under joint estimation is much smaller than under single-equation estimation, verifying joint estimation can provide more precise assessment of the mean WTP amounts than single-equation estimation. In fact, the joint estimation enables us to reject the null hypothesis that the two programs are equally valued, while single-equation variance estimates do not. This is an excellent example of how efficiency gains made possible by joint estimation could lead to different inferences.

[16] According to a simulation study of censored SUR models conducted by Riddel and Loomis, small efficiency gains are observed from joint estimation even when the regressor sets are identical across equations. The efficiency gains increase as the error correlation increases.



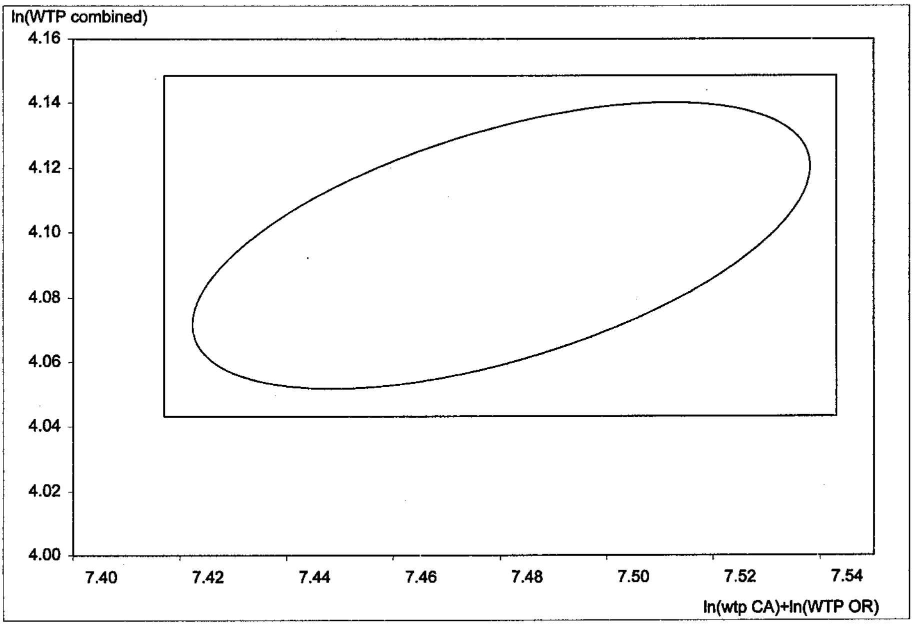
Note: The single-equation region is in the rectangle. The dotted line is the 45° line.

Figure 1. The 95% confidence ellipse for *E*[ln(*WTP* combined)] and *E*[ln(*WTP* California)]



Note: The single-equation region is in the rectangle. The dotted line is the 45° line.

Figure 2. The 95% confidence ellipse for *E*[ln(*WTP* combined)] and *E*[ln(*WTP* Oregon)]



Note: The single-equation region is in the rectangle.

Figure 3. The 95% confidence ellipse for *E*[ln(*WTP* combined)] and *E*[ln(*WTP* California) + ln(*WTP* Oregon)]

Examination of figures 1 and 2 also reveals that more uncertainty surrounds WTP values for the Oregon program than for the California program. This is not surprising, because only California residents were surveyed. The results indicate California residents have a lower and less well-defined preference for the out-of-state environmental amenity. Presumably, California residents derive use and existence value from the California site. The Oregon site may provide less use value for California residents because the distance to the site precludes frequent visits.

Finally, joint estimation permits a test for the existence of substitution between the Oregon and California programs. Specifically, we test the null hypothesis that the sum of the ln(*WTP*) for the Oregon and California programs is equal to the ln(*WTP*) for the two programs provided simultaneously; in other words:

$$H_0\colon E[\ln(WTP\ \text{Oregon})] + E[\ln(WTP\ \text{California})] = E[\ln(WTP\ \text{combined})],$$
$$H_A\colon E[\ln(WTP\ \text{Oregon})] + E[\ln(WTP\ \text{California})] > E[\ln(WTP\ \text{combined})].$$

The null hypothesis asserts that when the sum of the stated WTP amounts for the California and Oregon programs are close to the WTP for dual provision, substitution effects are negligible. Alternatively, the sum of the stated WTP for the individually provided program values may be greater than that for the dual program, indicating individuals substitute between similar programs.

The 95% confidence region for the substitution test is illustrated in figure 3. Because the region does not contain the 45° line, the null hypothesis of no substitution effects may be rejected. The data suggest stated WTP amounts do, in this survey, include substitution between programs.

Conclusions

This study introduces a simulated maximum-likelihood method for estimating mean WTP for several nonmarket goods from dichotomous-choice referendum data. Joint estimation is preferred to single-equation estimation because it affords smaller confidence regions around expected WTP values. Further, when WTP values are estimated jointly, the covariance of the WTP values may also be estimated, allowing for accurate joint tests of WTP across goods. This feature of joint estimation is especially useful when researchers are interested in testing for substitution between goods, or if they are interested in establishing a preference ordering among nonmarket goods. Gains from joint estimation will increase as the correlation among the WTP values for the goods increases.

We introduce an empirical model for estimating correlated WTP values using the GHK probability simulator and simulated maximum-likelihood techniques. The empirical model is applied to data from a survey designed to estimate WTP for fire reduction in spotted owl habitat in California and Oregon. Joint estimation leads to substantial reductions in the estimated variance of individual WTP amounts. The example illustrates tests for preference ordering among programs and substitution between programs. Using such tests, a researcher is able to clearly delineate the preferred program. The tests give evidence of substitution between potentially competing programs.

[*Received November 2000; final revision received October 2001.*]

References

Alberini, A. "Efficiency vs. Bias in Willingness-to-Pay Estimates: Bivariate and Interval Data Models." *J. Environ. Econ. and Manage.* 29(1995):169–80.

Borsch-Supan, A., and V. A. Hajivassiliou. "Smooth Unbiased Multivariate Probability Simulators for Maximum Likelihood Estimation of Limited Dependent Variable Models." *J. Econometrics* 58(1993): 347–68.

Breslaw, J. A. "Evaluation of Multivariate Normal Probability Integrals Using a Low Variance Simulator." *Rev. Econ. and Statis.* 76,4(1994):673–82.

Cameron, T. "A New Paradigm for Valuing Non-Market Goods Using Referendum Data: Maximum Likelihood Estimation by Censored Logistic Regression." *J. Environ. Econ. and Manage.* 15(1988): 355–79.

Cameron, T., and J. Quiggin. "Estimation Using Contingent Valuation Data from a 'Dichotomous Choice with Follow-up' Questionnaire." *J. Environ. Econ. and Manage.* 27(1994):218–34.

Carson, R. T. "Constructed Markets." In *Measuring the Demand for Environmental Quality*, eds., J. B. Braden and C. D. Kolstad. New York: Elsevier Science Publishers B.V. (North-Holland), 1991.

Carson, R. T., and R. C. Mitchell. "Sequencing and Nesting in Contingent Valuation Surveys." *J. Environ. Econ. and Manage.* 28(1995):155–73.

Geweke, J. "Bayesian Inference in Econometric Models Using Monte Carlo Integration." *Econometrica* 57(1989):1317–40.

Geweke, J., M. Keane, and D. Runkle. "Alternative Computational Approaches to Inference in the Multinomial Probit Model." *Rev. Econ. and Statis.* 76,4(1994):609–32.

Greene, W. H. *Econometric Analysis*, 4th ed. New York: Macmillan Publishing Co., 2000.

Hajivassiliou, V. "Smooth Simulation Estimation of Panel Data LDV Models." Work. Pap., Dept. of Econ., Yale University, 1990.

Hanemann, W. M., and B. Kanninen. "The Statistical Analysis of Discrete CV Data." In *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the U.S., E.U., and Developing Countries*, eds., I. Bateman and K. G. Willis. New York/London: Oxford University Press, 1999.

Hoehn, J. P. "Valuing the Multidimensional Impacts of Environmental Policy." *Amer. J. Agr. Econ.* 73(1991):289–99.

Hoehn, J. P., and J. B. Loomis. "Substitution Effects in the Valuation of Multiple Programs." *J. Environ. Econ. and Manage.* 25(1993):56–75.

Johnson, R. A., and D. W. Wichern. *Applied Multivariate Analysis*, 3rd ed. Englewood Cliffs NJ: Prentice-Hall Publishers, 1992.

Keane, M. "A Computationally Practical Simulation Estimator for Panel Data." *Econometrica* 62(1994): 95–116.

Lerman, S., and C. Manski. "On the Use of Simulated Frequencies to Approximate Choice Probabilities." In *Structural Analysis of Discrete Data with Applications*, eds., C. Manski and D. McFadden, pp. 305–19. Cambridge MA: The MIT Press, 1981.

Metropolis, N., and S. Ulam. "The Monte Carlo Method." *J. Amer. Statis. Assoc.* 44(1949):335–41.

Poe, G. L., M. P. Welsh, and P. A. Champ. "Measuring Differences in E(WTP) When Dichotomous Choice Contingent Valuation Responses Are Not Independent." *Land Econ.* 73(1997):255–67.

Riddel, M., and J. B. Loomis. "Joint Estimation of Multiple CVM Scenarios Under a Double-Bounded Questioning Format." *Environ. and Resour. Econ.* 12(1998):77–98.

Robert, C. P., and G. Casella. *Monte Carlo Statistical Methods.* New York: Springer Texts in Statistics, 1999.

Stern, S. "A Method for Smoothing Simulated Moments of Discrete Probabilities in Multinomial Probit Models." *Econometrica* 60(1992):943–52.

Appendix A: Specification of the SML Function for the Two-Good, Double-Bounded Model

We assume a two-good, double-bounded model, with correlated errors normally distributed as $N(\mathbf{0}, \Sigma)$. The corresponding log-likelihood for respondent i is:

$$
\begin{aligned}
\mathrm{LogL}_i = {} & I_i^{1111} * \log[1 + \Phi(z_i up, v_i up) - \Phi(z_i up) - \Phi(v_i up)] \\
& + I_i^{1110} * \log[\Phi(z_i up, v_i init) - \Phi(z_i up, v_i up) - \Phi(v_i init) + \Phi(v_i up)] \\
& + I_i^{1101} * \log[\Phi(z_i up, v_i low) - \Phi(z_i up, v_i mid) - \Phi(v_i low) + \Phi(v_i init)] \\
& + I_i^{1100} * \log[\Phi(z_i up) - \Phi(z_i up, v_i low) + \Phi(v_i low)] \\
& + I_i^{1011} * \log[\Phi(z_i mid, v_i up) - \Phi(z_i up, v_i up) + \Phi(z_i up) - \Phi(z_i init)] \\
& + I_i^{1010} * \log[\Phi(z_i up, v_i up) + \Phi(z_i init, v_i init) - \Phi(z_i init, v_i up) - \Phi(z_i up, v_i init)] \\
& + I_i^{1001} * \log[\Phi(z_i up, v_i init) + \Phi(z_i init, v_i low) - \Phi(z_i init, v_i mid) - \Phi(z_i up, v_i low)] \\
& + I_i^{1000} * \log[\Phi(z_i up, v_i low) + \Phi(z_i init) - \Phi(z_i init, v_i low) - \Phi(z_i up)] \\
& + I_i^{0111} * \log[\Phi(z_i init) + \Phi(z_i low, v_i up) + \Phi(z_i low) - \Phi(z_i init, v_i up)] \\
& + I_i^{0110} * \log[\Phi(z_i init, v_i up) + \Phi(z_i low, v_i init) - \Phi(z_i low, v_i up) - \Phi(z_i init, v_i init)] \\
& + I_i^{0101} * \log[\Phi(z_i init, v_i init) + \Phi(z_i low, v_i low) - \Phi(z_i low, v_i init) - \Phi(z_i init, v_i low)] \\
& + I_i^{0100} * \log[\Phi(z_i init, v_i low) + \Phi(z_i low) - \Phi(z_i low, v_i low) - \Phi(z_i init)] \\
& + I_i^{0011} * \log[\Phi(z_i low) + \Phi(v_i up) - \Phi(z_i low, v_i up)] \\
& + I_i^{0010} * \log[\Phi(z_i low, v_i up) + \Phi(v_i init) - \Phi(v_i up) - \Phi(z_i low, v_i init)] \\
& + I_i^{0001} * \log[\Phi(z_i low, v_i init) + \Phi(v_i low) - \Phi(v_i init) - \Phi(z_i low, v_i low)] \\
& + I_i^{0000} * \log[\Phi(z_i low, v_i low) - \Phi(v_i low) - \Phi(z_i low)],
\end{aligned}
$$

where $z_i init = (b_i init - \mathbf{b}_1'\mathbf{X}_{1i})/\sigma_1$, $z_i up = (b_i up - \mathbf{b}_1'\mathbf{X}_{1i})/\sigma_1$, $z_i low = (b_i low - \mathbf{b}_1'\mathbf{X}_{1i})/\sigma_1$, and $b_i init$, $b_i up$, and $b_i low$ are the initial, step-up, and step-down bids offered, respectively, to the ith respondent for good 1; $v_i mid = (c_i init - \mathbf{b}_2'\mathbf{X}_{2i})/\sigma_2$, $v_i up = (c_i up - \mathbf{b}_2'\mathbf{X}_{2i})/\sigma_2$, $v_i low = (c_i low - \mathbf{b}_2'\mathbf{X}_{2i})/\sigma_2$, and $c_i init$, $c_i up$, and $c_i low$ are the initial, step-up, and step-down bids offered, respectively, to the ith respondent for good 2.

Appendix B: Derivation of Confidence Regions for Expected WTP for Two Goods

Including the off-diagonal element in the error variance-covariance matrix leads to a rotation of the joint confidence ellipsoid for the two estimates and correct estimation of the major and minor axes. This provides a smaller confidence region, and a more accurate assessment of the variances of $E(WTP)$. The following discussion should make this clear.

Given that $\mathbf{b}$ is distributed asymptotically as $Np(\mathbf{b}, \mathbf{cov[b]})$, the expected WTP vector for the ith individual, $\mathbf{E(WTP}_i) = \mathbf{b'X}$, is distributed asymptotically as $N_k(\mathbf{b'X}, \mathbf{Xcov[b]X'})$. We can calculate a $100(1-\alpha)\%$ simultaneous confidence ellipsoid for the K different WTP values from (1) using the following inequality from Johnson and Wichern:

$$n(\mathbf{b'X} - \mathbf{b'X})'(\mathbf{Xcov[b]X'})^{-1}(\mathbf{b'X} - \mathbf{b'X}) <= k(n-1)/\left((n-k-p) * F_{k,n-k-p}(\alpha)\right),$$

where n = number of observations; p = number of regressors, including the constant term, in the kth equation; and K = the number of equations in the system.

The ellipse is centered around the estimated mean values for WTP. The lengths and slopes of the major and minor axes are determined by the eigenvalues and eigenvectors of $\mathbf{Xcov[b]X'}$. The lengths of the axes are equal to

$$\sqrt{\lambda_i}\,\sqrt{F(a)_{k,n-k-p}k(n-1)/n(n-k-p)},$$

where λ_i is the ith eigenvalue from $\mathbf{Xcov[b]X'}$. The slopes (directions) of the axes are given by the eigenvectors of $\mathbf{Xcov[b]X'}$.

When we assume $\mathbf{Xcov[b]X'}$ is a diagonal matrix, the major and minor axes correspond to the variable axes in Cartesian space. When the variances are the same across questions, the region becomes a k-dimensional hypersphere. If the off-diagonal elements are positive, then the region is flattened in the direction of the axis corresponding to the relatively smaller variance and the axes are rotated according to the level of correlation. We recall the formula for the area of an ellipse where $Area = \pi ac$, with a and c corresponding to the semi-axis lengths. As the correlation becomes stronger, the area of the confidence interval becomes smaller relative to one developed without regard to the correlation structure.