Estimating Willingness-To-Pay Using a Polychotomous Choice Function: An Application to Pork Products with Environmental Attributes

Sean P. Hurley

Douglas J. Miller

James B. Kliebenstein[1]

May 15, 2004

Selected paper prepared for presentation at the American Agricultural Economics Association

Annual Meeting, Denver, August 1-4, 2004

Abstract

This paper utilizes a polychotomous choice function to investigate the relationship between socioeconomic characteristics and willingness-to-pay for embedded environmental attributes. Specifically, a two-stage estimation procedure with an ordered probit selection rule is used to predict the premium payers and the magnitude of the premium they are willing to pay.

---

[1] Dr. Hurley is an Assistant Professor of Agribusiness at California Polytechnic State University, Dr. Miller is an Assistant Professor at Purdue University, and Dr. Kliebenstein is a Professor of Economics at Iowa State University. The authors gratefully acknowledge The Leopold Center for Sustainable Agriculture for its generous support of this work. For questions or comments regarding this paper, please email Dr. Hurley at: shurley@calpoly.edu.

Introduction

This paper analyzes observed consumer willingness-to-pay (WTP) for pork products with embedded environmental attributes. The data used in this study were collected from second-price sealed-bid auctions where the products used to elicit bids were two-pound packages of uniformly cut 1¼ inch boneless pork loin chops. Participants simultaneously bid on ten different packages of pork chops, each having differing environmental attributes. One package was labeled a "typical package" with no assigned environmental attributes, while the other nine packages contained pork produced in ways that embedded varying levels of environmental attributes. The attributes considered included ground water, surface water, and odors (air quality). Combinations of ground and surface water along with odor were also evaluated.

This paper utilizes econometric techniques to investigate the relationship between willingness-to-pay for embedded environmental attributes and socioeconomic characteristics of the respondents. There are three objectives in this paper. The first objective is to present an econometric model developed by Lee (1983) to accommodate data that has anchoring points within the distribution of the data.[2] The second objective is to predict the premium payers using socioeconomic characteristics that are currently used in the WTP literature. The third objective is to predict the magnitude of the WTP for the premium payers using the same variables that were used to predict the identity of the premium payers.[3]

Study Design and Data

Data were collected using a second-priced sealed-bid auction segmented into five bidding rounds. This auction was conducted using boneless pork loin chops defined as being from hogs

---

[2] An anchoring point for the purposes of this paper is defined as a point that has probability greater than zero, i.e., a point *a* within a continuous distribution such that Prob(x = *a*) > 0.
[3] Since there are a small number of non-premium payers who were negatively affected by the information, no attempt will be made to predict the magnitude of their WTP. A larger sample size would be needed for this task.

raised in farm production systems with varying environmental attributes. In the first three rounds, participants bid only on the physical attributes of the product having no other information except for the previous round's bids. In the fourth round, the participants were informed of the specific environmental attributes associated with the respective products.[4] In the fifth round, the implications of the embedded environmental attributes were further explained and the participants were allowed to bid a final time. Following Fox et al. (1995, 1996), wealth effects were controlled by randomly choosing one round and one product from that selected round to be the product sold.[5]

Two-pound packages of uniformly cut, boneless, 1¼ inch pork loin chops were used to elicit bids. These packages were made to look as uniform as possible to ensure bid responses only reflected the value of the environmental attribute. Participants simultaneously bid on ten different packages of pork chops each having different embedded environmental attributes. The packages were arranged in a row and placed on ice in one of three white coolers. Each package was labeled as "Package i", where i = 1,.., 10. After the third round, participants were informed that one package was a "typical package" with no particular environmental attributes. In this same round, participants were told that the other nine packages were from hogs produced under varying levels of environmental attributes pertaining to ground water, surface water, and odor reduction. Hog production with reduced odor was presented at two levels: a low level of 30-40 percent reduction, and a high level of 80-90 percent reduction over the "typical" production system.[6] Ground water and surface water impacts of the hog production system were also at two

[4] This release of information provides a means to determine the impact of releasing environmental information on participants' bids.
[5] Wealth effects may occur when participants change their bids because they won an earlier trial (Fox et al., 1995). See Davis and Holt for a discussion of wealth effects in experimental markets.
[6] The attribute of odor reduction was related to the production facility and its relationship to air quality. It was not related to the aroma of the pork chop, i.e., the product attribute was not proposing pork chops with different odors.

levels: a low level of 15-25 percent reduction and a high level of 40-50 percent reduction over the “typical” production system. Packages were provided with single attributes (only air, ground water, or surface water), double attributes, or all three embedded attributes. The double and triple attribute pork packages were all at the high reduction levels.

Experiments were conducted in six different areas of the United States: Ames, Iowa; Iowa Falls, Iowa; Manhattan, Kansas; Raleigh, North Carolina; Burlington, Vermont; and Corvallis, Oregon. Table A1 in the appendix gives the number of participants at each site. Three experiments were conducted at each site. Each experiment lasted approximately two hours and was conducted at 9:00 a.m., 11:30 a.m., or 2:00 p.m. To control for bias in package labeling, corresponding package number were switched with the assigned environmental attribute for each of the different time slots. A random sample of individuals from the area being studied was used to obtain participants for the study. This sample was obtained by a random generated computer sample drawn from telephone numbers in the respective local telephone directory. Each participant was paid forty dollars at the beginning of the experiment for their participation.

Table A2 in the appendix summarizes the changes in average bids from round three, the no information round, to round four, the environmental information round. The difference between the average high and low bid in the no-information third round is $0.35. This should reflect the difference in participant perception of the visual quality of the packages and does not represent a significant difference. For the entire group, the average bid increase for the two-pound package of pork loin chops with the highest level of embedded attributes was $0.94, while the bid for the typical package decreased by $0.52.

Methods Used to Model WTP Data

Many econometric methods have been used to analyze the relationship between WTP and socioeconomic characteristics. Menkhaus et al. (1992) and Melton et al. (1996a) used ordinary least squares (OLS), while Roosen et al. (1998) and Fox (1994) used more advanced models incorporating a two-stage analysis. Roosen et al. (1998) used Cragg's (1971) double hurdle model to investigate the relationship between WTP for apples with reduced pesticide use and socioeconomic characteristics. Fox (1994) used a Heckman (1976, 1979) two-stage procedure to evaluate WTP for milk with no trace of bovine somatotropin and socioeconomic characteristics.

There are two reasons Roosen et al. (1998) and Fox (1994) use more advanced modeling techniques over OLS. The first is associated with the method they used to collect their data. In both of their studies, they used a second-price sealed-bid multi-round auction for collecting WTP for food safety attributes. In their experiments, they initially endowed each participant with a product. Using the auction, they then asked the participants to bid on a product with food safety attributes. This bid reflected the participant's WTP to upgrade from their initial endowment to a product that had higher food safety attributes. Since Fox and Roosen et al. assumed that the product being bid on was no worse than the initial endowment, they placed a lower limit on the bids of zero. The information they collected was the WTP for the attribute. Hence, a censoring or limiting point at zero is induced for those participants who did not want to upgrade. The drawback to using OLS for censored data of this sort stems from the qualitative difference between the limit bids and the positive bids (Fox 1994). In this case, OLS is a biased estimator because it ignores the self-selection by the participants.

The second reason to use more advanced two-stage techniques is related to the nature of how consumers make decisions. Fox notes that "even in the absence of selection bias, the two stage method facilitates an intuitively appealing decomposition of the bidding decision (1994, p. 133)." By setting the lower limit for bids at zero, Roosen et al. (1998) and Fox (1994) caused the participants to self-select themselves into groups—those who want to pay a premium and those who do not. This implies that the modeling techniques they use needed to incorporate an aspect of self-selection. Standard OLS analysis cannot accommodate this in a one-stage procedure.

In contrast, the method used in this study for collecting WTP information elicits unbounded continuous values. In particular, the WTP measure was calculated from the change in bids from round three to round four which was not restricted to a lower or upper bound. Hence it would first appear that OLS estimation may be appropriate and advanced modeling techniques may not be necessary.

Table 1 provides the distribution of bids for the most environmental product. Upon examination of this table, there appears to be an issue that makes OLS inappropriate for analysis. This issue stems from approximately thirty-percent of the bids being zero. While the method of data collection allowed for an unbounded distribution of bids, the nature of the information given caused a discrete cluster point within the range of bids. In typical censored data applications such as Fox (1994) and Roosen et al. (1998), censored data has an upper and/or lower bound on the distribution. The data collected from this experiment has a discrete mass point within the distribution at zero. Hence using the OLS method to model this data will cause a bias in the estimates because the point zero will be weighted too heavily by standard estimation methods.

Table 1: Distribution of Willingness-to-Pay for the Most Environmental Product—The Product with High Ground Water, Surface Water, and Odor Improvements

	Premium Level (Interval) per Package							
	Below $0.00	**$0.00**	**$0.01-$0.49**	**$0.50-$0.99**	**$1.00-$1.49**	**$1.50-$1.99**	**$2.00-$2.49**	**Over $2.50**
Percent of Participants	7.6%	30.4%	9.7%	12.8%	13.1%	7.0%	8.2%	11.2%

Fox (1994) and Roosen et al. (1998) handled the issue of censoring by using a two-stage method for estimating the relationship between the dependent and independent variables. Since the models they use are very similar, only the method by Fox will be described. Fox (1994) employs a Heckman (1976, 1979) two-stage procedure to handle the censoring problem in his data. Heckman's approach considers the bias that arises to be a case of a specification error or a missing data problem. To handle this bias, he estimates the missing variable in the first stage, and then includes the estimates of the regressors in the second-stage. His method provides a measure of the degree of self-selection (Fox 1994). Fox explains that one of the advantages of this method is that it allows different variables to influence each decision, and it allows a single variable to have different effects for different groups.

Fox estimates the following equations:

(1) $Y_{1i} = X_{1i}\beta_1 + U_{1i} \qquad i \in G'$,

(2) $Y_{2i} = X_{2i}\beta_2 + U_{2i} \qquad i \in G$,

where G' is the subset of participants with non-zero bids. He notes that Equation 1 can be viewed as an inverse demand equation and Equation 2 is a choice function where Y_{2i} is a qualitative variable that takes on the value one when the participants pays a premium and zero otherwise.

If U_{1i} and U_{2i} are independent from each other, and U_{1i} has a conditional expectation of zero, then OLS can be used to estimate Equation 1. These error terms are usually not independent when self-selection occurs. Fox reports that the OLS estimator for the parameters in Equation 1 is typically biased. To account for the bias, he estimates the following equation:

(3) $$E(Y_{1i} | X_{1i}, U_{2i} \geq -X_{2i}\beta_2) = X_{1i}\beta_1 + (\lambda_i)\frac{\sigma_{12}}{(\sigma_{22})^{1/2}} ,$$

where σ_{12} and σ_{22} represent the covariance between U_{1i} and U_{2i} and the variance of U_{2i}, respectively. The term λ_i is defined to be the inverse Mill's Ratio.[7] Fox assumed that the joint distribution of U_{1i} and U_{2i} is bivariate normal.

To estimate this model, Fox (1994) employs Heckman's two-stage procedure. He first estimates Equation 2 as a probit equation on the full sample to obtain the probability that the bid will be positive. From this, he estimates the inverse Mill's Ratio for each observation. Finally, he estimates Equation 3 by OLS for the subset of participants who bid a positive amount. The OLS estimator of this final equation is consistent for β_1.

While useful for standard censored data with a lower bound, the two-stage methods that Roosen et al. (1998) and Fox (1994) used are not totally appropriate for modeling the WTP data from this study. The double hurdle method and the two-stage Heckman method are inappropriate for the data because the censoring in this study rests within the distribution rather than being a lower or upper bound. Also, their methods allow for only two choices. In contrast, the data from this study has three choices.

Lee (1983) offers a way of modeling this type of data using a two-stage procedure similar to the Heckman (1976, 1979) and double hurdle models. He suggests using a two-stage procedure that incorporates an initial polychotomous choice function, e.g., multinomial probit, in the first stage to estimate the discrete dependent variables. In the second stage, standard OLS procedures can be used to estimate the continuous dependent variables with the discrete variables factored out. One of the advantages of using Lee's model is that it can account for more than two choices in the selection process, whereas, the double hurdle model and the two-stage Heckman procedure cannot.

[7] See Fox (1994) for the calculation of the inverse Mill's Ratio.

Lee's Polychotomous Choice Selectivity Model

The model Lee proposes for handling dependent variables with mixed discrete and continuous variables can be set up as follows (1983). Suppose there is a polychotomous choice model with M categories and M regression equations. These equations can be written as:

(4) $y_s = x_s \beta_s + \sigma_s u_s$

(5) $y^*_s = z_s \gamma_s + \eta_s$ $(s = 1, \ldots, M)$,

where x_s and z_s are both exogenous explanatory variables. In Equation 4, σ_s is the standard deviation for a non-standardized distribution.[8] Equation 5 can be viewed as the choice equation, whereas, Equation 4 is the observed dependent variable when category s is chosen. Lee assumes the error terms, u_s and η_s, each have mean zero given the explanatory variables x_s and z_s for all s. All of the error terms in Equation 4 are assumed to have completely specified absolutely continuous marginal distributions.

Lee's model assumes that the dependent variables y_s are observed if and only if category s is chosen (1983). The choice of category s follows the rule

(6) $y_s^* > \max_{j=1,\ldots,M} y_j^* \quad where\, j \neq s.$

Letting the polychotomous variable K take on the values 1 to M, variable K takes the value of s if category s is chosen. Hence Equation 6 would imply that

(7) $K = s \quad \text{iff} \quad z_s \gamma_s > \varepsilon_s$

where

(8) $\varepsilon_s \equiv \max_{j=1,\ldots,M} y_j^* - \eta_s \quad where\, j \neq s.$

For each pair (u_s, ε_s), Lee defines the marginal distribution of u_s as $G_s(u)$ and the marginal distribution of ε_s as $F_s(\varepsilon)$. He states that by using the translation method, a bivariate distribution

[8] Note that this is equal to one when u_s is normally distributed.

of (u_s, ε_s) can be specified.[9] By letting $g_s(\cdot)$ be the density function of $G_s(\cdot)$, and defining the dummy variable D_s such that

(9) $D_s = 1$ iff $K = s$,

for s = 1, …,M, the log likelihood function can be specified. This function for a polychotomous choice model with random sample of size N can be written as

$$(10)\qquad \ln L = \sum_{i=1}^{N}\sum_{s=1}^{M}\left\{\begin{array}{l} D_{si}\ln g_s((y_{si} - x_{si}\beta_s)/\sigma_s) - D_{si}\ln\sigma_s \\ \quad + D_{si}\ln\Phi\Big((J_{1s}(z_{si}\gamma_s) - \rho_s J_{2s}(y_{si} - x_{si}\beta_s))/(1-\rho_s^2)^{1/2}\Big)\end{array}\right\}$$

where J_{1s} is the inverse of the cumulative distribution evaluated at $F_s(\cdot)$ and J_{2s} is the inverse of the cumulative distribution evaluated at $G_s(\cdot)$. By assuming that $\gamma_s = \gamma$, i.e., the set of explanatory variables across choices are the same for all s, and the marginal distribution of u_s are normally distributed N(0,1), a two-stage method can be used to estimate the equations

$$(11)\qquad y_s = x_s\beta_s - \sigma_s\rho_s\phi(J_{1s}(z_s\gamma))/F_s(z_s\gamma) + \eta_s \qquad (s = 1, \ldots, M),$$

where $\phi(\cdot)$ is the standard normal distribution function and the expectation of η_s given that choice s is selected equals zero. The conditional variance of η_s given that choice s is chosen is

$$(12)\qquad \begin{array}{l}\mathrm{var}(\eta_s | s\ is\ chosen) = \sigma_s^2 - (\sigma_s\rho_s)^2\big[J_{1s}(z_s\gamma) + \phi(J_{1s}(z_s\gamma))/F_s(z_s\gamma)\big] \\ \qquad\qquad\qquad\qquad\qquad * \phi(J_{1s}(z_s\gamma))/F_s(z_s\gamma)\end{array}$$

The estimator of this variance should be corrected for heteroscedasticity because the errors are correlated across sample observations.

There are two main reasons why Lee's model is the appropriate way to model the data in this study. First, due to the nature of the attribute that is being valued, there is a definite anchoring point within the distributions of bids. As mentioned above, this anchoring point causes a discrete mass point within a continuous distribution. The model by Lee is general enough to handle this issue by estimating the discrete variables first. Once these discrete

[9] Note that ρ_s can be defined as the correlation between u_s and ε_s.

variables have been estimated, they can be factored out leaving a continuous distribution with the appropriate probability structure, i.e., no discrete points with a positive probability mass.

The second reason Lee's model is appropriate is because it is intuitively appealing to think of the assessment function as a separate stage to developing a WTP measure. Thus in the first stage, the participant assesses the effect of the released information. In the second stage, the participant chooses the magnitude of the effect. Since this WTP was calculated from the difference between a naïve bidding round and a round with information, there arises a subjective classification of how the information affects the participant. This can be viewed as the assessment function from the behavioral model from Hurley and Kliebenstein (2003a). Hence, the derivation of WTP from the participant's standpoint can be viewed as a two-stage procedure similar to the double hurdle model where there is self-selection. The participants first decide what effect the information had on them, and then they choose the intensity of the effect. This decision causes a self-selection process that also can be handled by Lee's generalized model.

Lee's model is general enough to allow different explanatory variables for determining the magnitude of each category. Hence the explanatory variables used to explain the magnitude of the WTP for the premium payers can be different from the explanatory variables for the negative premium payers.

Two-Stage Estimation with an Ordered Probit Selection Rule

Information shocks pertaining to product attributes can have a natural self-selection aspect to them. When maximizing consumers are given new information on a product, they must decide on how that new information impacts their purchase decision. They decide whether the information has a positive, neutral, or negative effect. In this sense, the consumers can be viewed as self-selecting themselves into a group. Once they have decided which group they belong, they can reallocate their resources to maximize their utilities. Since this self-selection process has a natural ordering to it, an appropriate selection rule would be an ordered probit rule that has three choices—a negative premium, no premium, and a positive premium.

Let z equal the ex post categorical realization of whether the consumer was negatively affected, denoted by a zero, not affected, denoted by a one, or positively affected, denoted by a two. The ordered probit part of the model can be written as:

(13) $z^* = \alpha' W + u$

where, $z = 0$ if $z^* < 0$, i.e., the participant is negatively affected by the information;

1 if $0 \leq z^* \leq \mu_1$, i.e., the participant is not affected by the information;

2 if $z^* > \mu_1$, i.e., the participant is positively affected by the information.

Equation 13 can be considered a latent utility function where z* is the unobserved utility. The term z is the observed choice that is made by the consumer. It is assumed that the error term u is distributed as standard normal. The term μ_1 is an unknown threshold parameter that is estimated with the explanatory values. The matrix W is a set of explanatory variables and the vector α is the set of corresponding coefficients. While Lee's model can account for the explanatory variables being different for each category, it is also assumed that the explanatory variables for the ordered probit model are the same for each category. The WTP equation can be written as:

(14) $WTP_s = \beta_s' X_s + \varepsilon_s$,

where s represents one of the three categories chosen—premium payers, negative premium payers, or those unaffected. WTP_s is the WTP vector of the subset of participants that fall into category s. The term ε_s is assumed to be normally distributed with mean zero, has a standard deviation of σ_s, and has a correlation of ρ_s with u from the ordered probit model. The matrix of explanatory variables, X_s, includes LAMBDA, which is the estimated bias that occurs due to the self-selection process. The corresponding coefficient vector, β_s, is the vector of explanatory variables.

To estimate this model, Greene (1995) describes this two-stage procedure as having four steps in the process. The first step is to estimate the ordered probit equation using maximum likelihood estimation on all the observations, which accounts for the discrete variable. The second step is to select the subset of observations to use in the OLS regression. The third step is to estimate this equation by OLS including the correction term that takes into account the choice that was selected. The final step is to correct the asymptotic covariance matrix for the estimates of this subset of observations. The econometric software LIMDEP was used to estimate this model.

Empirical Results

It is assumed that the explanatory variables are the same for Equation 13 and 14. The estimated model has two WTP equations with a trichotomous choice function to be estimated. Equation 13 is estimated first. The bias component from the self-selection process is estimated for each participant and then used as a regressor in the corresponding OLS estimation. Then Equation 14 is estimated for the positive premium payers.[10] The group whose WTP was zero does not need to be estimated by the OLS procedure because this group has been estimated using the ordered probit.

The explanatory variables for both equations are a subset of the socioeconomic characteristics and derived variables collected from the experiment. The explanatory variables related to socioeconomic characteristics are taken from the literature on WTP for attributes. Specifically, the papers by Roosen et al. (1998), Menkhaus et al. (1992), and Melton et al. (1996a) are the major sources of the socioeconomic factors that enter Equations 13 and 14.

[10] Due to the small number of negative premium payers, this group will not be estimated.

These explanatory variables are described in Table A3 in the appendix. These are participant's age, household income, participant's education, and participant's gender. For this model, location of the experiment is also used as a variable. Both pork consumption and number of people living in the household are used in this model for both definitions. These variables include continuous, discrete, and categorical variables.

The first equation estimated is the ordered probit equation. The explanatory variables used in these ordered probit equations are a constant term and all of the explanatory variables in Table A3 excluding EDU1, EDU2, INC1, INC2, and LOC7.[11] In this case, the first two responses in education and income and the location of the second experiment done in Raleigh are being used as the bases of comparison for their respective categories. Roosen et al. (1998), Menkhaus et al. (1992), and Melton et al. (1996a) are used to hypothesize the sign of the explanatory coefficients.

There are three multi-response categories used in this model. The first two are education and income. It is hypothesized that a higher education level will increase the probability of the participant being a premium payer. It is also expected that the coefficients increase in magnitude as the education level goes up. Like education, income is also hypothesized as positive and having higher coefficients for higher income levels.

The other categorical variable in this model is related to location of the experiment. Since there is nothing in the literature which gives an a priori expectation to the effect a location can have on WTP, a benefit hypothesis will be investigated. Within this benefit hypothesis, it is expected that locations closer to high concentrations of hog production will tend to have a higher benefit received from consuming pork with embedded environmental attributes. It was stated

[11] Due to the extremely small number of participants falling into EDU1 and INC1, EDU2 and INC2 were also excluded to avoid collinearity between the constant term and the income and education category.

above that the second experiment in Raleigh is used as the basis for location. It is expected that the location variable associated with Iowa Falls will have a positive effect on the probability of WTP. The cities of Manhattan and Corvallis are expected to have a negative coefficient because they are farther away from the high concentrations of hog production compared to Raleigh. Hence, these areas would receive less benefit than Raleigh would. It is unclear what sign Ames would have based on the second Raleigh experiment.[12] It is expected that the first Raleigh experiment should not add significantly to the probability of being a premium payer relative to the second Raleigh experiment, i.e., the coefficient is expected to be close to zero.

Monthly pork consumption of the participant, PORKM, the number of people living in the participant's household, NOINHOUS, and the participant's age, AGE, are hypothesized to have a negative effect. The final variable that is standard in the literature is the participant's gender, GENDER. Taking from the findings of Andreoni and Vesterlund (2001), it is hypothesized that women will have a higher probability of paying a premium.

Table 2 provides the result of the ordered probit model. Three estimated parameters were significant at the five or ten percent level of significance. The constant term and the estimated threshold parameter were significant at the five-percent level. Gender was significant at the ten-percent level and had the expected positive sign. This implies that being a woman increased the likelihood of being a premium payer. All of the other estimated variables were not significant. The variables for education have signs consistent with the a priori expectations, i.e., positive sign.[13] This implies that a person who had at least a high school diploma has a higher probability of being a premium payer. While the sign was consistent with expectations, the magnitudes of the effect were not. It was hypothesized that the magnitude of the effect would increase as

[12] Ames like Raleigh is situated geographically close to pork production facilities, but neither could be considered in the heart of pork production like Iowa Falls is.

[13] All of these education levels are being compared to the group of participants with less than a high school degree.

education level increased. This is not the case. A participant with a Bachelors degree had the highest magnitude effect for being a premium payer. A participant with a Doctorate degree has the second highest probability of being a premium payer, while a person with some college has the third highest magnitude effect. The group of participants that had the lowest magnitude effect was the group that has some technical, trade, or business schooling.

Table 2: Ordered Probit Estimates for the Ex Post Categorical Realization of Whether the Participant Was Negatively Affected, Not Affected, or Positively Affected [a]

Variable	Coefficient[b]	Standard Error	Mean of Variable
Constant	1.2780*	0.6138	
NOINHOUS	0.0076	0.0485	2.6869
PORKM	-0.0113	0.0150	5.8290
GENDER	0.2443**	0.1502	0.5988
AGE	-0.0052	0.0049	47.7362
LOC1	0.0609	0.2763	0.1489
LOC2	0.2136	0.2716	0.1824
LOC3	-0.0079	0.2911	0.0942
LOC4	-0.2573	0.3030	0.0821
LOC5	0.0691	0.2764	0.1763
LOC6	0.1422	0.2660	0.1824
INC3	-0.2859	0.2620	0.1376
INC4	0.1669	0.2544	0.1865
INC5	0.0851	0.2614	0.1407
INC6	0.3906	0.3334	0.1040
INC7	0.0780	0.3180	0.0703
INC8	-0.2289	0.3309	0.0599
INC9	-0.0184	0.4273	0.0398
INC10	-0.1795	0.3265	0.0734
EDU3	0.2925	0.4754	0.1220
EDU4	0.0831	0.4792	0.0854
EDU5	0.3063	0.4439	0.2530
EDU6	0.3873	0.4668	0.2409
EDU7	0.1871	0.5056	0.0732
EDU8	0.2939	0.4694	0.1220
EDU9	0.3326	0.5416	0.0579
Threshold parameter for index			
μ_1	1.1847*	0.1168	
N = 329			

(a) A premium payer is a participant who increased her bid for the most environmental package from round three to round four.

(b) An asterisk * implies that the coefficient is significant at the five-percent level of significance and a double asterisk ** implies significance at the ten-percent level.

Excluding income and location, two other variables have consistent signs, while one does not. The other variables that were not significant but had consistent signs were age and number of times pork is consumed in a month. Both of these variables had a negative effect on the probability of being a premium payer. Hence a participant who was older had a lesser probability of being a premium payer. Also, the probability that a participant was a premium payer decreases as he/she consumes more pork in a month. The variable that had an inconsistent sign and was insignificant was number in household. It was hypothesized that this variable would have a negative effect. This variable took on a positive and very small value.

Some of the income variables had positive signs as expected, while others were inconsistent with expectations. The basis of comparison for the income levels was the category with income less than or equal to $20,000. The variables for the income levels from $30,000 to $70,000 all have the expected positive coefficient. While this group of variables has consistent signs, they do not have the hypothesized increasing magnitudes. This implies that if the participant fell in one of these income categories, he/she would have a higher likelihood of being a premium payer compared to someone who makes $20,000 or less. The income variables for the income levels over $70,000 have the inconsistent sign of being negative. Hence, having a high income implies that the participant was less likely to be a premium payer compared to someone who makes $20,000 or less. The participants who fell in the income range of $20,000 to $30,000 were also less likely to be premium payers compared to those participants who made less than $20,000.

All of the location variables have insignificant signs. Some of the variables have consistent signs with the benefits hypothesis, while others do not. Iowa Falls has the expected positive sign, while Burlington has the expected negative sign. The first Raleigh experiment had a negative coefficient, but it is extremely close to zero as expected. Manhattan and Corvallis were expected to have a negative coefficient. Both of these variables had the unexpected positive sign. This would imply that the benefits hypothesis may not be enough to explain how location affects WTP for environmental pork. Although the sign for Ames was a priori indeterminate, the estimated coefficient is positive. This implies that participants in Ames are more likely to be premium payers compared to participants from Raleigh.

Table 3 provides the frequencies of actual and predicted outcomes for participant group placement from the estimated ordered probit equation for each definition of WTP. The columns show the predicted outcomes from the model, while the rows show the actual outcomes from the data. The major result is that the probit equation failed to predict which participants were negatively affected by the environmental information. The model also had difficulty predicting who was not affected by the environmental information.

Table 3: Frequencies of Actual and Predicted Outcomes from the Estimated Ordered Probit for Definition 1 of Willingness-to-Pay[a]

	Predicted Outcome			
Actual Outcome	Negatively Affected	Not Affected	Positively Affected	**Total**
Negatively Affected	0	4	21	25
Not Affected	0	8	92	100
Positively Affected	0	6	198	204
Total	0	18	311	329

(a) A premium payer is a participant who had a higher bid for the most environmental package compared to the typical package within round four.

The probit equation had a strong tendency to predict premium payers over the other groups. Of the 329 participants, the equation picked 311 premium payers. Of this group, ninety-two participants were not actually affected by the information and twenty-one participants were negatively affected. The probit equation was not able to predict any negative premium payers correctly. Furthermore, the model had trouble predicting the participants who were not affected by the environmental information. This probit equation does not do a very good job predicting the three different categories using the core variables used in the WTP literature.

Table 4 presents the conditional OLS model predictions of the premium magnitude for those who were affected positively by the environmental information. In the second column, the estimated standard errors without the heteroscedasticity correction are presented, while in the third column, the estimated standard errors corrected for heteroscedasticity are presented. The explanatory variables used to predict the magnitude for this group are assumed to be the same as the variables used to predict which category each participant falls into. The predicted signs and magnitudes for this equation will be the same as the first-stage probit parameters. Hence, it is expected that income and education will have positive signs with increasing magnitudes. The number in household, monthly pork consumption, and age are all expected to have negative coefficients. Gender is expected to have a positive coefficient. The location variables are also expected to have the same signs as in the probit equation. Also included with these explanatory variables is LAMBDA, which is an adjustment factor for the bias caused by the clustering of zeros.

Table 4 shows that the value for number in household, age, gender, and monthly pork consumption all have signs consistent with a priori expectations. Age has the expected negative coefficient and is significant at the five-percent significance level. At the ten-percent significance level, both gender and monthly pork consumption are significant. Gender has the expected positive coefficient, while monthly pork consumption has a negative coefficient. While the value for the household parameter is not significant, it has the expected negative sign.

Table 4: Second-Stage OLS Analysis of the Positive Premium Payers for Definition 1 of Willingness-to-Pay[a]

Variable	Coefficient[b]	Standard Error (Uncorrected)	Standard Error (Corrected)	Mean of Variable
Constant	-5.2814	6.1650	4.9218	
NOINHOUS	-0.0201	0.0924	0.0713	2.7598
PORKM	-0.0755**	0.0577	0.0458	5.6193
GENDER	1.6749**	1.1205	0.9156	0.6324
AGE	-0.0567*	0.0255	0.0230	46.8369
LOC1	0.5133	0.5235	0.5429	0.1471
LOC2	0.9499	1.0290	0.8407	0.1961
LOC3	-0.6417	0.4547	0.4226	0.0931
LOC4	-1.3752	1.3421	1.1100	0.0735
LOC5	0.6058	0.5299	0.5265	0.1716
LOC6	0.9225	0.7621	0.6748	0.1863
INC3	-2.5784*	1.4503	1.2601	0.1141
INC4	0.2129	0.8331	0.6922	0.2028
INC5	-0.3956	0.5728	0.4428	0.1484
INC6	1.2828	1.6751	1.4142	0.1285
INC7	-0.3034	0.6158	0.6836	0.0791
INC8	-2.2129*	1.1993	0.9553	0.0495
INC9	-0.7742	0.6465	0.6357	0.0396
INC10	-1.7473*	1.0130	0.8748	0.0644
EDU3	2.6061*	1.6032	1.2314	0.1225
EDU4	0.7413	0.8735	0.5234	0.0784
EDU5	2.5661*	1.6343	1.2564	0.2500
EDU6	2.8897**	1.9599	1.5745	0.2647
EDU7	3.5634*	1.2595	1.1795	0.0686
EDU8	2.8889*	1.6236	1.2557	0.1324
EDU9	2.9007*	1.7731	1.4013	0.0539
LAMBDA	10.9237**	8.2374	6.7337	0.5898
N	204			
R^2	0.2041			
Log-Likelihood	-355.0125			
Log-Likelihood (Restricted)	-378.2970			

(a) A premium payer is a participant who had a higher bid for the most environmental package compared to the typical package within round four.

(b) An asterisk * implies that the coefficient is significant at the five-percent level of significance and a double asterisk ** implies significance at the ten-percent level.

When examining the category of education, there are many coefficients that are significant at either the five or ten-percent level of significance. The only education variable that is not significant is the one pertaining to some technical, trade, or business schooling. At the ten-percent level of significance, the variable related to a Bachelors degree is significant. For all of the other education levels, the parameters are significant at the five-percent level of significance. The magnitudes of the education coefficients indicate that the higher education levels tend to have higher magnitudes over the lower education levels.

Similar to the probit equations above, the variables for income in the OLS model tend to not have the expected signs. In Table 4, there are only two income levels that have the expected positive sign. These are the income level associated with $30,000 to $40,000 and the income level associated with $50,000 to $60,000. The rest of the income variables are negative. There are three income levels that are significantly negative at the five-percent level of significance—$20,000 to $30,000, $70,000 to $80,000, and the highest income category.

The location variables are not significant at either the five or ten-percent level of significance. Among these variables, only two have the hypothesized sign. Burlington has the expected negative coefficient, while Iowa Falls has the expected positive coefficient. Manhattan and Corvallis have the unexpected sign of positive. Ames has a positive coefficient, while the first Raleigh experiment has a negative coefficient.

The bias adjustment coefficient LAMBDA shows the level of bias due to the zeros has a positive and significant effect at the ten-percent level of significance. Hence, deleting the zeros and running OLS on the remaining observations would cause a serious bias to occur in the estimates on the coefficients. Using a likelihood ratio test, the null hypothesis that all coefficients are zero for this model can be rejected at the five-percent level of significance. The critical value for this test at the five-percent level of significance is 38.89, while the calculated likelihood ratio from the model is 46.56. Hence, the variables in this model do have explanatory power.

Summary and Conclusions

Bid data for pork chops with embedded environmental attributes were analyzed to determine which consumers would pay a premium and how much they would pay. A premium payer was defined as a participant who had a positive willingness-to-pay for pork produced under the most environment-friendly system. It was found that approximately sixty-two percent of the participants in the study had a positive WTP for the most environmental package of pork. Approximately thirty percent of the participants had no WTP, i.e., they bid zero for the environmental improvements, while approximately eight percent had a negative WTP.

A model by Lee was presented to demonstrate a method for handling discrete mass points, i.e., anchoring points, within a continuous distribution. Lee's model uses a two-stage procedure that incorporates an initial polychotomous choice function in the first stage to estimate the discrete dependent variables. In the second stage, OLS procedures were used to estimate the continuous dependent variables with the discrete variables factored out.

There are two advantages to modeling the data from this study using Lee's model. By using this method, participants can be classified as premium or non-premium payers and the magnitude of the premium can be predicted. From a marketing point of view, an important task is to predict premium payers from non-premium payers so marketing efforts can be focused on targeted consumers. From the research standpoint, there is another advantage to using Lee's model. Since his model can account for anchoring points within a distribution, economic experiments are no longer confined to truncating WTP values, i.e., researchers no longer have to design experiments that assume information impacts have no adverse effects. This model allows researchers more flexibility when initially designing their experiments.

This paper utilized the standard variables used in the WTP literature coupled with Lee's model to predict who were premium payers and non-premium payers using an ordered probit equation. It was found that this equation did not predict very well the three different categories using the core variables used in the WTP literature. The only significant variables in the equation were gender and the constant term. Education, income, monthly pork consumption, number of people living in the household, and age all had insignificant effects. This implies that the standard variables in the WTP literature are not sufficient to separate who was positively, negatively, and not affected by the environmental information released.

Once the ordered probit equation was estimated, OLS procedures were used to predict the magnitude of the positive premiums utilizing the standard WTP variables from the literature. Gender had a significant positive effect on premiums, while monthly pork consumption had a significant negative effect. Age had a significantly negative impact on the premium. The value for the household parameter had an insignificant negative effect on a participant's premium. Many education coefficients had a significant effect on the premium—higher education levels tended to have higher premium effects over the lower education levels. Variables for income tended to not have the expected impacts on premiums. Two income levels positively affected premiums—income levels associated with $30,000 to $40,000 and $50,000 to $60,000. The rest had unexpected negative impacts. Three income levels that were significantly negative were $20,000 to $30,000, $70,000 to $80,000, and the highest income category. Location variables did not have a significant effect on premiums. The bias adjustment coefficient LAMBDA showed that the level of bias due to the zeros had a significant and positive effect. Hence, the bias from the anchoring point of zero is an important factor that needed to be factored into the OLS estimation procedure.

References

Andreoni, James and Lise Vesterlund. "Which Is The Fair Sex? Gender Differences in Altruism." *The Quarterly Journal of Economics* 116 (February 2001), 293-312.

Cragg, John G. "Some Statistical Models for Limited Dependent Variables with Application to Demand for Durable Goods." *Econometrica* 39 (September 1971): 829–844.

Davis, Douglas D. and Charles A. Holt. *Experimental Economics*. Princeton, New Jersey: Princeton University Press, 1993.

Fox, John A. "Essays in the Measurement of Consumer Preferences in Experimental Auction Markets." Ph.D. Dissertation, Department of Economics, Iowa State University, 1994.

Fox, John A., Brian L. Buhr, Jason F. Shogren, James B. Kliebenstein, and Dermot J. Hayes. "A Comparison of Preferences for Pork Sandwiches Produced from Animals with and without Somatotropin Administration." *Journal of Animal Science* 73 (1995): 1048-1054.

Fox, John A., Dermot J. Hayes, James B. Kliebenstein, and Jason F. Shogren. "Consumer Acceptability of Milk from Cows Treated with Bovine Somatotropin." *Journal of Dairy Science* 77 (1994): 703-707.

Fox, John A., Dermot J. Hayes, Jason F. Shogren, and James B. Kliebenstein. "Experimental Methods in Consumer Preference Studies." *Journal of Food Distribution Research* (July 1996): 1-7.

Fox, John A., Jason F. Shogren, Dermot Hayes, and James B. Kliebenstein. "CVM-X: Calibrating Contingent Values with Experimental Auction Markets." *American Journal of Agricultural Economics* 80 (August 1998): 455-465.

Freund, John E. *Mathematical Statistics*. Fifth Edition, Englewoods Cliff, New Jersey: Prentice Hall, 1992.

Greene, William H. *LIMDEP version 7*.0 User's Manual. Bellport, NY: Econometric Software Inc., 1995.

Hoffman, Elizabeth, Dale J. Menkhaus, Dipankar Chakravarti, Ray A. Field, and Glen D. Whipple. "Using Laboratory Experimental Auctions in Marketing Research: A Case Study of New Packaging for Fresh Beef." *Marketing Science* 12 (Summer 1993): 313-338.

Heckman, James J. "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* (1976): 475-492.

Heckman, James J. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (January 1979): 153–161.

Hurley, Sean P. and James B. Kliebenstein. "A Tale of Two Premiums—Examining Bids From a Multiple Round Vickrey Auction with Differing Information Sets." Selected paper presented at the American Agricultural Economics Association Annual Meeting, Montreal, Canada, July 27-30, 2003a.

Hurley, Sean P. and James B. Kliebenstein. "Interpreting Bids From a Vickrey Auction when There Are Public Good Attributes." Selected paper presented at the American Agricultural Economics Association Annual Meeting, Montreal, Canada, July 27-30, 2003b.

Hurley, Sean P. and James B. Kliebenstein. "Determining the Benefits of Environmental Improvements in Pork Production and Their Sustainability: A Community-Based Study of Iowa's Pork Industry." A report prepared for the Leopold Center for Sustainable Agriculture: Ames, IA, November 1998.

Lee, Lung-Fei. "Generalized Econometric Models with Selectivity." *Econometrica* 51 (March 1983): 507–512.

Maddala, G. S. "*Limited Dependent and Qualitative Variables in Econometrics*. Cambridge, England: Cambridge University Press, 1983.

Melton, Bryan E., Wallace E. Huffman, Jason F. Shogren, and John A Fox. "Consumer Preferences for Fresh Food Items with Multiple Quality Attributes: Evidence from an Experimental Auction of Pork Chops." *American Journal of Agricultural Economics* 78 (November 1996): 916-923.

Melton, Bryan E., Wallace E. Huffman, and Jason F. Shogren. "Economic Values of Pork Attributes: Hedonic Price Analysis of Experimental Auction Data." *Review of Agricultural Economics* 18 (October 1996): 613-627.

Menkhaus, Dale J., George W. Borden, Glen D. Whipple, Elizabeth Hoffman, and Ray A. Field. "An Emperical Application of Laboratory Experimental Auctions in Marketing Research." *Journal of Agricultural and Resource Economics* 17 (July 1992): 44-55.

Roosen, Jutta, John A. Fox, David A. Henessey, and Alan Schreiber. "Consumer's Valuation of Insecticide Use Restrictions: An Application to Apples" *Journal of Agricultural And Resource Economics* 23 (December 1998): 367-384.

Trost, Robert P. and Lung-Fei Lee. Technical Training and Earnings: A Polychotomous Choice Model With Selectivity." *The Review of Economics and Statistics* 66 (February 1994), 151-156.

Appendix

Table A1: Number of Participants by Area

Experiment Area	Number of Participants
All areas	329
Ames, IA	49
Manhattan, KS	60
Raleigh, NC (6/28/97)	31
Burlington, VT	27
Iowa Falls, IA	58
Corvallis, OR	60
Raleigh, NC (6/27/98)	44

Table A2: Participant Bid Levels by Environmental Attribute Information (All Participants)

Pork Chop Environmental Attributes (Level of Improvement over Typical)	Average Bid Level per Package ($) No Information	Environmental Attribute Added	Premium Bid Absolute Change*
No Particular Environmental Attributes (Typical)	4.13	3.61	**-0.52**[a]
Odor 30-40%	4.26	3.87	**-0.39**[a]
Odor 80-90%	4.05	3.92	*-0.13*[b]
Ground water 15-25%	3.91	3.85	-0.06[b,c]
Ground water 40-50%	4.03	3.94	-0.09[b,c,d]
Surface Water 15-25%	4.15	3.99	*-0.16*[b,c,d]
Surface Water 40-50%	4.06	4.10	0.04[b,c,d]
Odor 80-90%/Ground Water 40-50%	4.25	4.56	**0.31**[e]
Odor 80-90%/Surface Water 40-50%	4.17	4.58	**0.41**[e]
Odor 80-90%/Ground Water 40-50%/Surface Water 40-50%	4.19	5.13	**0.94**

*Corresponding letters indicate that at the five percent level of significance the null hypothesis of the two bid changes were equal could not be rejected. Also, note that the bold and italic changes represent a significant difference from zero at the 0.001 and 0.05 level respectively.

Table A3: Variable Description for Each Estimated Equation

Variable	Description
NOINHOUS	Number of people living in the household
PORKM	Number of times per month pork is consumed by participant
GENDER	1 if female, 0 otherwise
AGE	Age of the participant
LOC1	1 for experiments conducted in Ames, IA; 0 otherwise
LOC2	1 for experiments conducted in Manhattan, KS; 0 otherwise
LOC3	1 for experiments conducted in Raleigh, NC in 1997; 0 otherwise
LOC4	1 for experiments conducted in Burlington, VT; 0 otherwise
LOC5	1 for experiments conducted in Iowa Falls, IA; 0 otherwise
LOC6	1 for experiments conducted in Corvallis, OR; 0 otherwise
LOC7	1 for experiments conducted in Raleigh, NC in 1998; 0 otherwise
INC1	1 if household income is less than $10,000; 0 otherwise
INC2	1 if household income is between $10,000 and $20,000; 0 otherwise
INC3	1 if household income is between $20,000 and $30,000; 0 otherwise
INC4	1 if household income is between $30,000 and $40,000; 0 otherwise
INC5	1 if household income is between $40,000 and $50,000; 0 otherwise
INC6	1 if household income is between $50,000 and $60,000; 0 otherwise
INC7	1 if household income is between $60,000 and $70,000; 0 otherwise
INC8	1 if household income is between $70,000 and $80,000 ; 0 otherwise
INC9	1 if household income is between $80,000 and $90,000; 0 otherwise
INC10	1 if household income is over $90,000; 0 otherwise
EDU1	1 if highest level of education achieved was eight grade
EDU2	1 if highest level of education achieved was eleventh grade
EDU3	1 if highest level of education achieved was high school or G.E.D.
EDU4	1 if highest level of education achieved was some technical, trade, or business school
EDU5	1 if highest level of education achieved was some college, no degree
EDU6	1 if highest level of education achieved was a Bachelors degree
EDU7	1 if highest level of education achieved was some graduate work, no degree
EDU8	1 if highest level of education achieved was Masters degree
EDU9	1 if highest level of education achieved was a Doctorate degree