Comparison of Approaches to Estimating Demand for Payment for Environmental Services

Monthien Satimanon[*] and Frank Lupi[+]

*Department of Economics, Michigan State University, East Lansing, MI 48824
+Department of Agricultural Food and Resource Economics, Michigan State University, East Lansing, MI 48824

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association 2010 AAEA,CAES, & WAEA Joint Annual Meeting, Denver, Colorado, July 25-27, 2010 (10496)

Comparison of Approaches to Estimating Demand for Payment for Environmental Services

1. Introduction

This paper proposes a comparison of both parametric and semiparametric estimation of willingness to pay (WTP) for environmental services. Payment for environmental services (PES) is an approach that uses economic incentives either provided by public or private sector to protect natural resources. PES programs range from classical soil and water conservation to the new areas such as drinking and farming water supply and carbon sequestration. Hence, PES programs have been of recent interest globally and have led to an increasing number empirical studies. Two important questions for PES studies are what determines the willingness to pay (WTP) or demand for PES? and what determines participation in PES programs by payment recipients?. Both of these questions have been answered by estimating the dichotomous choice (binary choice) models by using standard Probit or Logit estimation. The standard procedure of this contingent valuation can be found in the work of Haneman (1984) and Haab and McConnell (2003). In this binary response valuation models, WTP usually refers to conditional mean that is derived from estimated parameters under given underlying distributional assumption. The problem with this set up is that the welfare evaluations will crucially depend on these specific distributions. Unlike the linear probability model, the consistencies of estimated parameters depend on the underlying distribution as well as the conditional variance of the estimated model. In this context, semiparametric estimation provides an interesting alternative since it allows flexible functional form for conditional variance.

Semiparametric methods have been used in estimation of binary choice model for a long period of time, as summarized in Li and Racine (2008). In most theoretical studies, the semiparametric models have been compared with parametric binary choice model by simulation. Horowitz (1992) found that semiparametric models will be more robust when the estimated model contains heteroskedasticity. Klein and Spady (1993) and Li (1996) also found strong support for the semiparametric model. In empirical application of semiparametric methods to welfare measurement in binary choice model, Chen and Randall (1997), Creel and Loomis(1997), An(2000), Cooper(2002), and Huang et.al(2009) found out that the semipametric results are robust and can be used as a complementary procedure along with the parametric estimation. Also, it can be used to check whether the parametric model encounters any inconsistency problems because of underlying distribution, unobserved heterogeneity, and heteroskedasticity.

The methods that will in this paper to compare estimation in the binary choice models are Probit (Probit), Klein and Spady estimator (KSE), Heteroskedasticy Probit (HP), and sieve semiparametric estimator (S). The comparison includes the estimated parameters as well as the estimated standard errors since the WTP is derived from these parameters.

The data used for the comparison of welfare measures comes from a study of the demand for payment for environmental services (PES) in eastern Costa Rica. The data set come from the extended surveys of Ortega-Pacheco et.al. (2009). The respondents are asked to vote "Yes" or "No" in the response to additional payment for the people who live in the upstream and mountainous area to preserve the quality as well as quality of water sources that will be used in the lower area. The bid value has been provided in standard referendum contingent valuation. The goods here are clearly defined since the people who live along the downstream self-financed their existing water supply and already pay the water fees monthly. With the new estimation methods and extended data from previous study, the results show that the choice of model can influence the results.

2. Binary Response Model and Estimation Methods

The estimated model in this study is specified as random utility model with a linear utility function as in Haab and McConnell(2003). The change in deterministic utility of the proposed contingent valuation is

$$v_{1j} - v_{0j} = (\alpha_1 - \alpha_0)z_j + \beta_1(m_j - t_j) - \beta_0 m_j \quad (1)$$

where $v_{1j} - v_{0j}$ is change in indirect utility after the respondents are asked to choose the required payment for the program, m_j is income, t_j is the payment asked, z_j is an m-dimensional vector of exogenous variables related to individual j. By assuming that the marginal utility of income is constant between two states of contingent valuation then, the probability of voting can be defined as adding one error terms to the difference equation (1). Thus the probability of voting "Yes" for each respondent j become

$$\Pr(y_j = yes) = \Pr(\alpha z_j - \beta t_j + \varepsilon_j > 0) \quad (2)$$

where $\varepsilon_j \equiv \varepsilon_{1j} - \varepsilon_{0j}$ or the error in valuation of the difference in (1). The conventional process in estimating the parameters of the model in (2) is to specifying the distribution of the error terms. In most of the studies, ε_j are independently and identically distributed (IID) with mean zero. Then, either the underlying distribution of normal and logistic will be used as in the case of Probit and Logit estimation. For comparison of this study only the basic Probit will be used.

2.1 Probit

In, Probit model, the probability of "Yes" will be model in term of latent variable that $y_j = 1$ if $y_j^* = \alpha z_j - \beta t_j + \varepsilon_j > 0$ and probability of "No" will be defined as $y_j = 0$ if $y_j^* = \alpha z_j - \beta t_j + \varepsilon_j \leq 0$. Or, binary response model is in the form of index function $y_j = 1\{\alpha z_j - \beta t_j + \varepsilon_j > 0\}$.

Also, for the errors term, it is assumed to be $\varepsilon_j \sim f(z_j + t_j) \sim N(0, \sigma^2)$ and $\theta = \varepsilon / \sigma \sim N(0,1)$. Hence the distribution is assumed to be as followed.

$$\begin{aligned} &\Pr(\varepsilon_j < \alpha z_j - \beta t_j) = \Pr(\theta < \frac{\alpha z_j}{\sigma} - \frac{\beta}{\sigma} t_j) \\ &= \Phi(\frac{\alpha z_j}{\sigma} - \frac{\beta}{\sigma} t_j) \end{aligned} \tag{3}$$

where $\Phi(x)$ is the cumulative standard normal distribution. Then, the parameters can be estimated up to a scale as well as the marginal effects. In order to estimate this model, the maximum likelihood estimation will be used. Defining a new 1 x (m+1) parametric vector $\mathrm{B} = \{\alpha / \sigma, -\beta / \alpha\}$ where m+1 is the dimension of covariates including constant terms, and define the data vector $X = \{z_j, t_j\}$, the likelihood function will be

$$L(B \mid y, X) = \prod_{j=1}^{T} \left\{ \left[\Phi\left(\frac{X'B}{\sigma} \right) \right]^{y_j} \left[1 - \Phi\left(\frac{X'B}{\sigma} \right) \right]^{1-y_j} \right\} \tag{4}$$

Then, the familiar log likelihood function is

$$\ln L(B \mid y, X) = \sum_{j=1}^{T} \left\{ y_j \ln\left[\Phi\left(\frac{X'B}{\sigma} \right) \right] + (1 - y_j) \ln\left[1 - \Phi\left(\frac{X'B}{\sigma} \right) \right] \right\} \tag{5}$$

2.2 Heteroskedasticity Probit (HP)

The simple estimation will be modified with the unobserved heterogeneity by incorporating the heteroskedasticity into standard Probit model. The variance σ^2 will be varying as a function of independent variables. The variance will be a multiplicative function of z_j as followed

$$\sigma_j^2 = \{\exp(\gamma z_j)\}^2 \tag{6}$$

Substituting this variance into equation (3) yields multiplicative heteroskedastic probit model.

$$\Phi\left(\frac{\alpha z_j}{\exp(\gamma z_j)} - \frac{\beta}{\exp(\gamma z_j)} t_j\right) \tag{7}$$

Then, the log likelihood function will become

$$\ln L(B \mid y, X) = \sum_{j=1}^{T} \left\{ y_j \ln\left[\Phi\left(\frac{X'B}{\exp(\gamma z_j)} \right) \right] + (1 - y_j) \ln\left[1 - \Phi\left(\frac{X'B}{\exp(\gamma z_j)} \right) \right] \right\} \tag{8}$$

The result of estimation from equation (5) and (8) will be useful in positing whether our estimated model contain heteroskedasticity or not. Further more, the other assumptions that can be relaxed is functional specification of ε_j.

2.3 Klein and Spady (KS)

In Klein and Spady(1993), the distribution function of errors term $\Phi\left(\frac{X'B}{\sigma}\right)$ will be estimated rather than assumed to be normal distribution. However, one strong assumption that need to be put forth is the that ε_j and X are independent. Let define the estimated function to be $\hat{G}(X'\beta)$, then the true probabilities $G(X'\beta)$ can be written in terms of the density of $X'\beta$ on y, denoted $f(X'\beta \mid y)$. The specific form is as followed

$$G(X'\beta) = \frac{\Pr(y_j = 1) f(X'\beta \mid y = 1)}{\Pr(y_j = 1) f(X'\beta \mid y = 1) + (1 - P) f(X'\beta \mid y = 0)} \tag{9}$$

They proposed the way to estimate equation (9) by a leave-one-out nonparametric kernel estimator that is given by

$$\hat{G}_{-j}(X'\beta) \equiv \hat{E}_{-j}(y_j \mid X'\beta) = \frac{(nh)^{-1} \sum_{j=1, j \neq i}^{n} y_j K\left(\frac{X_j'\beta - X_i'\beta}{h}\right)}{(nh)^{-1} \sum_{j=1, j \neq i}^{n} K\left(\frac{X_j'\beta - X_i'\beta}{h}\right)} \tag{10}$$

where n is number of observations, h is bandwidth, and K is researchers' choice of kernel function. Klein and Spady suggested estimating the parameter β by maximum likelihood methods. The estimated log likelihood function is

$$\ln L(B \mid y, X) = \sum_{j=1}^{T} \left\{ y_j \ln\left[\hat{G}(X'\beta)\right] + (1 - y_j) \ln\left[1 - \hat{G}(X'\beta)\right] \right\} \tag{11}$$

Under some regularity conditions, the estimated parameters $\hat{\beta}_{KS}$ is $\sqrt{n}$ -consistent and has asymptotic normal distribution given by $\sqrt{n}(\hat{\beta}_{KS} - \beta) \to N(0, \Omega_{KS})$, where

$$\Omega_{KS} = \left[E\left\{ \frac{\partial \Pr(\varepsilon_j < X'\beta)}{\partial \beta} \left(\frac{\partial \Pr(\varepsilon_j < X'\beta)}{\partial \beta} \right)' \left[\frac{1}{\Pr(\varepsilon_j < X'\beta)(1 - \Pr(\varepsilon_j < X'\beta))} \right] \right\} \right]^{-1} \tag{12}$$

This estimation is semiparametrically efficient since it reaches the efficiency bound. It means this estimator is asymptotically as efficient as the nonlinear least squares(as well as linear least squares) estimator based on the known functional form of

$G(X'\beta)$. However, this semiparametric estimator is less efficient than the least squares when the true functional form of $G(X'\beta)$ is known since the estimation is two steps rather than one step estimation. One point that is worth noting is that the estimated parameters can not be directly compared to the P and HP model given different underlying distribution. Therefore, the average partial effects or average derivative estimates have to be calculated as stated as in Li and Racine(2008).

2.4 Sieve estimator, Probit model with distribution-free heteroskedasticity (S)

Sieve estimation refers to one class of semiparametric estimation that solves the problem of infinite dimensional parameter. The sieve method employs the optimization routine that tries to optimize the criterion function over finite approximated parameter spaces (sieves). The sieve method, in the simplest form, might be similar to how we choose the bandwidth and numbers in plotting the histogram. As pointed out by Chen (2007), the method of sieves is very flexible in estimating complicated semiparametric models with (or without) endogeneity and latent heterogeneity. It can easily incorporate prior information and constraints, and it can simultaneously estimate the parametric and nonparametric parts, typically with optimal convergence rates for both parts. Khan(2005) proposed a estimation method that is a further expansion of Horowitz(1992) method. The important assumption is the conditional median restriction to ensure the identification of estimated parameters β.

$$med(\varepsilon_j \mid X) = 0 \tag{13}$$

and symmetric distribution of the error terms the local nonlinear least squares estimator for $y_j = 1\{X'\beta + \varepsilon_j > 0\}$ is defined as

$$\hat{\beta} = \arg\min_{\beta\in\Theta x1} \frac{1}{n}\sum_{i=1}^{n}\left[y_i - \Phi\left(\frac{X'B}{h_n}\right)\right]^2 \tag{14}$$

where h_n is a sequence of positive numbers such that $h_n \to 0$ as $n \to \infty$. This estimator will yield the estimated β with one of the estimated element to be normalized to 1 as usual for semiparametric estimation. Blevins and Khan(2009) provides the procedure to estimation equation(14), they suggested the use of probit criterion function for the sieve nonlinear least squares. The criterion function is

$$Q_n(\theta, l) = -\frac{1}{n}\sum_{i=1}^{n}[y_i - \Phi(X'\beta \cdot \exp(l(X)))]^2 \tag{15}$$

where $l(X)$ is finite dimensional scaling parameter and $\beta = (\theta', 1)'$ is a finite vector of parameters. Then, they introduce a finite-dimensional approximation of $l(X)$ using a linear-in-parameters sieve estimator as in Chen(2007). They define the estimator as followed. Let $b_{oj}(x_i)$ denotes a sequence of known basis function and $b^{\kappa_n}(x_i) = (b_{oj}(x_i), ..., b_{o\kappa n}(x_i))'$ for some integer κ_n. Then, the function $\exp(l(X))$ will be estimated by the following sieve estimator $\exp(b^{\kappa_n}(x_i)'\lambda_n)$ where λ_n is a vector of constants. Let $\alpha_n \equiv (\Theta, \lambda_n) \in \mathrm{A}_n$ where A_n is the sieve space. The estimator can be defined by $\hat{\alpha}_n = \arg\min_{\alpha\in\mathrm{A}_n} \frac{1}{n}\sum_{i=1}^{n}\left[y_i - \Phi\left(X'\beta \cdot g_n(x_i)\right)\right]^2$. The choice of $g_n(x_i)$ is arbitrary and can be any possible series such as power and polynomial series, and spline. In this study, we estimate the $g_n(x_i)$ by exponential function that contains the power series of (x_i) as a domain. Chen(2007) proved that the estimated parameters from sieve estimation will be asymptotically normal and consistent. However, in this paper we are interested in the estimation of willingness to pay so we have to apply further step in estimation. From estimation of equation (15), we can get the estimation of $g_n(x_i)$, then we will plug this one in the probit estimation of equation (3). The main reason that we proceed in two step estimation is that we can apply the results from Ackerberg et.al.(2009) in order to estimate the asymptotic variance by using parametric approximation since it requires less computation power to get the variance of the estimate of β and willingness to pay. Then, the average partial effects as well as willingness to pay will be easily computed by the usual delta method, and Krinsky and Robb, respectively.

To conclude this section, there are certain insights that might be gained from comparing these four methods of estimation. The probit and heteroskedastic probit models are computationally simple and should be more efficient if the underlying distributions are correctly specified. On the other hands, the two semiparametric models in this paper are not nested with each corresponding probit and heteroskedastic probit, but heuristic comparison can be made as in Beluzzo(2004). Results of Probit, HP, and S can be compared to see whether the underlying normal distribution is a valid assumption or not. Also, results from Probit, HP and S can be compared to see whether the there is a problem of heteroskedasticity in the data generating process or not.

3. Data and Estimating Results

3.1 Data

The data in this study came from eastern Costa Rica. The research sites contain not only the two communities as in Ortega-Pacheco et al (2009) but also four communities (Table1) within the region that were recently surveyed. The communities' local water supply is too polluted for drinking water usage due to heavy use of chemical substances in nearby pineapple and banana plantation. Their drinking water supply comes from aqueducts that pipe in water from the forested upper reaches of their watershed. The communities have local water boards that oversee the construction and maintenance of these water systems and levy monthly fees for water. However, changes in land use in the upper reaches of the watersheds threaten the quality of the communities' drinking water. To protect their water, the communities are considering PES programs to keep land uses from changing in the upper watershed. The surveys assess local resident's willingness to pay to finance these PES programs. The payment vehicle is a monthly surcharge on their water bill. There are 1179 completed interviews from the surveys. The dependent variable is the binary choice variable of voting "Yes" or "No" for the program for a particular fee (cost) in addition to the current water bill. The independent variables are the fee (cost) of the program, female dummy, age, high school dummy, household income, and household characteristics.

Table 1. Communities in the study

Community	Herediana	Cairo-Francia	Florida	Alegría	Milano	Iberia
Interviews	397	164	248	131	136	103

In the Table 2, summary statistics of variables used in estimation are presented along with their description. In total, there are 1141 observations to be used after using respondent with reported income.

Table 2. Data description and descriptive statistics (N=1141)

Description	Measurement	Mean	Std. Dev.	Min	Max
Response to " Would you vote for or against the program if you would have to pay [cost] colones more on your water bill (yes or for = 1, no or against = 0)	1 Or 0	0.659	0.474	0	1
Monthly cost of program (on top of current water bill) from the vote question. Defined in the preamble to the vote question and varied across respondents	Colones	1243.087	758.098	400	2400
A dummy for sex of the respondent (female = 1 male = 0)	1 Or 0	0.725	0.447	0	1
Age of respondents		43.176	15.113	18	93
Number of household member under age 18		1.515	1.418	0	9
A dummy for schooling (high school or more = 1 otherwise = 0)	1 Or 0	0.120	0.326	0	1
Monthly household income	Colones	142364.4	141374.8	7000	2000000

The respondents are asked to Vote “Yes” or no for the proposed increase in the monthly water fee. From, the observations about 66 percent of people voted “Yes”. This variable will be the dependent variable y_j in the estimated model. The bid value for each respondent will range from 400 to 2400 colones, this represents the additional water fee that each respondent has to pay for the PES program. This additional fee is a direct payment to people who manage land upstream. The recipient of the fee payment will in return conserve the resources in the surrounding catchments. This will ensure the preservation of both water quality and quantity. The other dependent variables are female which indicates the sex of respondents, age of respondents, number of household members under age 18, and education level of respondent. Average monthly income is 142,364 colones that is slightly higher than national average household income of 140,000 colones and slightly lower than household incomes in urbanized and metropolitan areas of Costa Rica's Central Valley (Ortega-Pachego et al, 2009).

3.2 Estimation Results

The methods presented in Section 2 were estimated using Vote "yes" as dependent variable. The set of other covariates are monthly cost, female, age, number of member less than 18, and education. Table 3 gives coefficient estimate obtained from Probit, Heteroskedastic Probit, Klein and Spady, and sieve estimator.

Overall, the key variables in the model are significant and yield expected signs. The additional monthly cost has a negative impact on the probability of voting "Yes" in all four estimation methods. If the sex of respondent is female, then it will have lead to lower probability of voting yes to the PES. Furthermore, the age of respondents and number of household member under age 18 both have negative effects on the probability of voting for the program. For the education variable, if the respondent has high school degree, it might lead to higher chance of voting. However, only under the estimation by KSE methods, does education becomes significant. The monthly income also has a positive effect on probability of voting "yes".

Table 3. Estimated coefficients and mean willingness to pay

Dependent variable =1 if respondent vote "Yes", 0 if the vote is "No".	Probit	HP	KSE	S
Monthly Cost	-0.0004	-0.0007	-0.001	-0.0004
	(-8.29)	(-3.10)	(-3.04)	(-8.47)
Female	-0.242	-0.272	-0.624	-0.252
	(-2.53)	(-1.42)	(-2.68)	(-2.55)
Age	-0.015	-0.022	-0.357	-0.016
	(-4.85)	(-2.92)	(-2.37)	(-4.92)
Household members	-0.06	-0.028	-0.146	-0.086
	(-1.95)	(-0.33)	(-2.04)	(-2.53)
Education	0.096	0.890	0.443	0.055
	(0.65)	(0.94)	(2.28)	(0.38)
Income	2.56-e06	6.6e-06	5.46-e06	2.73e-06
	(4.95)	(1.89)	(3.05)	(5.49)
Intercept	1.568	1.990		1.639
	(7.02)	(3.58)		(7.13)

Note: 1) t-statistics is reported in parenthesis

In these methods of estimation, the individual coefficients estimated are not directly comparable. So, the use of average partial effects or the marginal effects at the expected value of overall distribution needed to be computed for comparison. These marginal effects are of interest because they inform whether on average what will be the effect of each variables on the probability of voting “Yes” for the program. In contingent valuation, they are the marginal effects of each variable on the probability of voting yes on the referendum when evaluate at the average or expected of the underlying distribution. They are presented in Table 4.

Table 4. Average partial effects.

Marginal Effects	Probit	HP	KSE	S
Monthly Cost (in 10000 colones)	-1.438	-1.7	-1.119	-1.532
	(-9.13)	(-7.11)		(-9.36)
Female	-0.077	-0.09	1	-0.082
	(-2.60)	(-2.62)		(-2.58)
Age	-0.005	-0.004	-0.003	-0.005
	(-5.01)	(-2.90)		(-5.08)
Household members	-0.02	-0.027	-0.021	-0.027
	(-1.96)	(-2.24)		(-2.55)
Education	0.03	0.042	-0.008	0.018
	(0.66)	(0.83)		(0.38)
Income (in 10000 colones)	0.008	0.011	0.008	0.008
	(6.17)	(5.12)		(6.71)
Willingness to Pay				
Mean	2340.69	2582.50	1250.15	2233.96

Note: 1) The average partial effects for Probit and S comes standard integration and delta methods.

2) The average partial effects of HP comes from MEHETPROB command in Stata.

3) The average partial effects of KSE are based on Li and Racine(2008) average derivatives and parameter of fem_1 needs to be normalized to 1.

4) Willingness to pay for Probit and HP are calculated based on Krinsky and Robb method

5) Willingness to pay for KSE is calculated by the method as in Beluzzo(2004)

It can be clearly seen that most of the partial effects from these models are quite similar in sign and magnitude. However, there are certain differences in education and monthly cost variables. For monthly cost, an increase in 1,000 colones of water fee will lead to the lower probability of voting "Yes" by 14-15 percent for probit and HP and S model while the KSE model will lead to the lowering of probability by 11.9 percent. For education, the estimated average partial effect of KSE is of the wrong sign and magnitude is a lot lower than the probit and HP model. One of explanation for the difference is that both Probit and HP yield the similar willingness to pay as presented in table 2, but KSE yield the lower willingness to pay.

In table 2, the average mean willingness to pay (WTP) is 2340 colones for Probit model, 2490 colones for HP model, only 1250 colones for KSE model, and 2236 colones for S model. For Probit, HP, and S model the estimated mean WTP are twice the current average monthly bills of 1015 colonses(Ortega-Pacheco et.al 2009), however, WTP from KSE model is of the same side as the current level of water fees. Hence, it is uncertain to say which WTPs are more appropriate to use; however, as usual, it is clear evidence that WTP in general are certainly depends more on the assumption of underlying distribution as well as conditional variance. The KSE method as well as Turnbull estimator that do not assume such specific form of estimation might be a good representative for the lower bound for the mean WTP. Moreover, in controlling for conditional heteroskedasticity, the WTP form the model with flexible functional form is lower than the parametric one by about 10 percent. It implies that we should employ both flexible functional form and standard heteroskedasticity probit in estimation.

4. Conclusion and further study

This study has presented a comparison of approaches of estimation for the willingness to pay in the contingent valuation set up. The standard linear random utility model has been estimated by Probit, heteroskedastitc probit, and semiparametric estimation. The estimation results come from the contingent valuation study of payment for environmental services in Costa Rica. The referendum of the study is asking respondent to vote “Yes” or “No” to an additional monthly water fee to pay for conservation of water resources by the group of people who live upstream. The results from the parametric estimation yield similar results as the previous study in term of WTP; however, the estimation from the semiparametric gives significantly lower estimation of WTP when there is relaxation of underlying distribution assumption. On the other hand, if the model allows only a flexible functional form of conditional variance(S), the estimated WTP is slightly lower by about 10 percent compared to the standard heteroskedasticity(HP) model.

Nevertheless, there is still more work to be done within this area of research. The current S estimation model still has no canned package that can be easily applied by empirical research. Secondly, regarding the difference between parametric and semiparametric estimation, further investigation on the effect of conditional variance on WTP needed using explorations by theoretical modeling and simulation. Thirdly, it is possible that the low WTP form KSE model might come as a result of bimodality as appeared in the Beluzzo (2004), and we need to further explore this issue and the use of better semiparametric estimators that can solve this issue might be useful. Also, the use of quantile regression might be of interest.

References

Ackerberg, Daniel, Chen, Xiahong, and Hanh Jinyong, 2009. A Practical Asymptotic Variance Estimator for Two-Step Semiparametric Estimators. Working Paper UCLA, Sep 2009.

An, Mark Yuying, 2000. A semiparametric distribution for willingness to pay and statistical inference with dichotomous choice contingent valuation data. *American Journal of Agricultural Economics* 82, 487–500.

Belluzzo Jr., Walter, 2004. Semiparametric approaches to welfare evaluations in binary response models. *Journal of Business & Economic Statistics* 22, 322–330.

Blevins, Jason, and Khan, Shakeeb, 2009. Distribution-Free Estimation of Heteroskedastic Binary Response in Stata, working paper Duke University.

Boyle, Kevin J., 2003. Contingent valuation in practice. In: Champ, P.A., Boyle, K.J., and Brown, T.C., *A Primer on Nonmarket Valuation*, Dordrecht: Kluwer.

Chen, Heng Z., and Randall, Alan, 1997. Semi-nonparametric estimation of binary response models with an application to natural resource valuation. *Journal of Econometrics* 76, 323–340.

Chen, Xiahong, 2007. Large Sample Sieve Estimation of Semi-nonparametric Models, chapter 76 in *Handbook of Econometrics*, *Vol. 6B*. North Holland: Elsevier.

Cooper, Joseph C., 2002. Flexible functional form estimation of willingness to pay using dichotomous choice data. *Journal of Environmental Economics and management* 43, 267–279.

Creel, Michael, and Loomis, John, 1997. Semi-nonparametric distribution-free dichotomous choice contingent valuation. *Journal of Environmental Economics and Management* 32, 341–358.

Haab, Timothy C., and McConnell, Kenneth E, 2003. *Valuing Environmental and Natural Resources*. Northampton: Edward Elgar.

Hanemann,W. Michael, 1984.Welfare evaluations in contingent valuation experiments with discrete responses. *American Journal of Agricultural Economics* 66, 332–341 (August).

Horowitz, Joel L., 1992. A smoothed maximum score estimator for the binary response model. *Econometrica* 60, 505–532 (May).

Huang, Ju-Chin, Nychka, Douglas W., and Smith, Kerry V.(2008), Semi-parametric discrete choice measures of willingness to pay, *Economics Letters* 101, 91-94.

Khan, S. 2006. Distribution Free Estimation of Heteroskedasticity Binary Response Models Using Probit Criterion Functions. Working Paper, Duke University.

Klein, Roger W., and Spady, Richard H., 1993. An efficient semiparametric estimator for binary response models. *Econometrica* 61, 387–421.

Li, C., 1996. Semiparametric estimation of the binary choice model for contingent valuation. *Land Economics* 72, 462–473.

Li, Q, and Racine, Scott L., 2008. *Nonparametric Econometrics.* Princeton: Princeton University Press.

Máñez Costa, M., and Zeller, M., 2005. Calculating incentives for watershed protection. A case study in Guatemala,in: M. Markussen, R. Buse, H. Garrelts, M. Máñez

Costa, S. Menzel, & R. Marggraf (Eds) *Valuation and Conservation of Biodiversity*, pp. 297–314 (Berlin: Springer).

Ortega-Pacheco, Daniel V., Lupi, Frank, and Kaplowitz, Michael. 2009, Payment for environmental services: estimating demand within a tropical watershed. *Journal of Natural Resources Policy Research* 1(2), 189-202.