Valuing Water Resources in Developing Countries: A Semiparametric Approach to Valuation Models

(Valuing Water Resources: A Semiparametric Approach)

Walter Belluzzo, Jr.
University of São Paulo at Ribeirão Preto

Abstract

Valuation of the benefits from the preservation of water resources is often of interest to policy makers and funding institutions. In developing countries, valuation studies are potentially useful for designing funding policies when inequality is a concern. To fulfil this goal, however, valuation studies must provide detailed information about the whole distribution of benefits, not only its mean. This article applies semiparametric methods to acquire that information and presents their application to a valuation study involving an important Brazilian River basin. Results obtained suggest that the semiparametric model reveal a heterogeneity structure that cannot be accommodated by the logistic model. Specifically, it was found that the willingness-to-pay distribution is bimodal. As the logit places mass symmetrically, it tends to overestimate net benefits, leading to the undue acceptance of the project.

Departamento de Economia
Av. dos Bandeirantes, 3900
14.040-900 Ribeirão Preto, SP
Brazil

email: belluzzo@usp.br
vox: +55 16 633 1018
fax: +55 16 633 4488

1 Introduction

Preservation of water resources is a critical aspect of the sustainable development agenda. Intense economic and demographic growth often lead to the deterioration of water resources in developing countries. In Brazil, specifically, several river basins have reached a considerably compromised situation. In face of the substantial costs associated with projects for recuperation and preservation of those river basins, policy makers and funding institutions often require detailed evaluation of the benefits accruing from them.

The problem in evaluating these benefits is the public good nature of the preservation of water resources. There is no market where individuals reveal their preferences. As a result, usual demand analysis is prevented and alternative approaches to elicit preferences are required. One of these approaches is the so-called contingent valuation method. Based on survey data, this method provides an estimate of the welfare change due to public policy projects, and has been used to value a wide variety of public goods, including water resources. One of the distinguishing characteristics of the contingent valuation method is use *referendum* questions, whose responses are modelled using binary response methods.

To date, most contingent valuation studies found in the literature use *parametric* binary response models. In addition to assumptions regarding the functional form of the conditional mean willingness-to-pay, these models also assume that the willingness-to-pay distribution belongs to some known parametric family, the normal and logistic distributions being the most popular choices. Even though imposing such a restrictive assumption is justifiable for computational easy, it is important to recognize that misspecification of the underlying willingness-to-pay distribution may lead to biased estimates, with clear effects on welfare analysis.

The misspecification of the underlying willingness-to-pay distribution is particularly important in the context of the water resources projects in developing countries. In general, evaluation of such projects focus on a representative individual and assumes that the associated tax price will be the same for all individuals. As a result, only estimates of the mean benefit are necessary. However, as it often happens, the tax price is either progressive or regressive and welfare analysis depends critically on the whole conditional distribution of benefits, not only its central tendency. Therefore, imposing a particular shape to the willingness-to-pay distribution may be particularly restrictive in these circumstances.

In this context, semiparametric models, which impose less strident restrictions on the underlying distribution, represent an interesting modeling alternative. Several distribution-free models are available for estimating binary response models: Manski (1975), Cosslett (1983), Klein & Spady (1993), Horowitz (1992), among others. Nonetheless, applications of these methods to valuation models, are not numerous. Apparently, only Creel & Loomis (1997), Chen & Randall (1997), and Li (1996) considered distribution-free methods for the estimation of valuation models.

The purpose of this article is to present a semiparametric modeling approach to valuation models. This approach consists of the application of the Klein and Spady's (1993) estimator and related methods. Specifically, because the intercept is not identified, in order to recover the valuation function it was approximated assuming that the random term has zero mean. Additionally, I show how welfare evaluations can be computed from the estimated model.

The methods proposed are illustrated with the valuation of a project for management and improvement of an important Brazilian river basin. The Doce river basin is located in southeast Brazil, with a total area of 83400 Km^2 spread over two states. As a result of the intense economic development observed in this region, especially in the so called steel valley with mining and steel metallurgy activities, the basin have been suffering a steady process of deterioration.

The project being valued involve investments intended to preserve the areas that still are in good condition and to recover those already compromised. It is closely related to federal legislation for the management of water resources in Brazil, which requires the creation of an administrative agency for each river basin. These agencies are responsible for determining and implementing investment plans and the cost share for all consumers, domestic and industrial.

The results obtained suggest that the willingness-to-pay distribution is bimodal, with important consequences to welfare evaluations. Specifically, it was found that net benefits are significantly overestimated by the logit, leading to the undue acceptance of the project according to the Kaldor-Hicks criterion.

2 Binary Response Valuation Model

In a typical contingent valuation study, each individual is presented with a single bid value, t, through a $referendum$ question like "would you be willing to pay \$t for the implementation of this project?" In general, the bid value is randomly drawn from a pool of less than 10 values.

Given this general framework, suppose that individuals derive utility from the non-market good whose provision is to be changed and from monetary income. Assume further that individuals reveal their true preferences through the *referendum* question[1]. Then we can express responses as the result of a process of utility maximization, so that a "yes" answer implies that

$$\Delta v(m, t, A; \theta) = v(1, m - t, A; \theta) - v(0, m, A; \theta) \geq 0 \tag{1}$$

where m stands for monetary income, A is a vector of individual's characteristics, θ is a vector of parameters, $v(1, m)$ is the indirect utility function when the project is implemented and $v(0, m)$ when it is not. For future reference, the function $\Delta v(\,\cdot\,; \theta)$ will be called the utility difference function.

Alternatively, we can consider the dual problem of expenditure minimization. In

[1] That means that all questions about incentive compatibility and biased answers discussed in the contingent valuation literature are conveniently resolved.

this case a “yes” answer implies that

$$s(m, A; \theta) = e(0, v(0, m, A; \theta)) - e(1, v(0, m, A; \theta)) \geq t \tag{2}$$

where $e(i, \cdot) = v^{-1}(i, \cdot)$, $i = 0, 1$, are the expenditure functions associated with each state of the good’s provision. For future reference the function $s(m, A; \theta)$ will be called the valuation function. Clearly, the deterministic versions of these approaches, i.e., without the introduction of random terms, give the same result by duality. However, as was shown by McConnell (1990), when random terms are introduced results are the same only in the case where certain conditions for the marginal utility of income are satisfied.

The construction of the econometric models for these approaches is based on the fact that the values assumed by the utility difference function and the valuation function are not directly observable. Instead, only an indicator y, covariates $x \equiv \{m, A\}$ and the bid values t are observed. Specifically, introducing additive random terms to (1) and (2) we have that

$$y = \begin{cases} 1 & \text{if } z(x, t; \theta) \geq \varepsilon \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where $z(x, t; \theta)$ is equal to $\Delta v(x, t; \theta)$ in the utility difference approach and equal to $s(x; \theta) - t$ in the valuation function approach.

Note that for $z(x, t; \theta) = \Delta v(x, t; \theta)$, the model (3) is equivalent to the well-known random utility model. This is the approach proposed by Hanemann (1984) for modeling contingent valuation data. The case where $z(x, t; \theta) = s(x; \theta) - t$ is the approach proposed by Cameron (1988) and Cameron & James (1987).

Provided that there is sufficient information about the distribution of ε, denoted by F_{ε}, the expected gains conditional on a vector of individual characteristics x and cost share c can be easily computed. Define gains as $G(x, c) \equiv s(x; \theta) - c$. Then, noting that the probability of a "yes" answer can be written as[2]

$$\int_{-\infty}^{\infty} \mathbb{1}(G(x,c) \geq 0)\, dF_\varepsilon,$$

the conditional expectation of positive gains (gains) is

$$G^+(x,c) = \int_{-\infty}^{\infty} G(x,c)\, \mathbb{1}(G(x,c) \geq 0)\, dF_\varepsilon, \tag{4}$$

where $\mathbb{1}(\cdot)$ represents the indicator function, which assumes the value 1 if the condition is satisfied and zero otherwise.

Substituting $G(x,c) < 0$ for the inequality in the indicator function (4) gives the conditional expectation of negative gains (losses), denoted by $G^-(x,c)$. Note that (4) can be used to obtain expected gains conditional on any vector of individual characteristics x and cost c. In practice, however, specifying and reporting all individual characteristics generally is not feasible. A better approach in the present case may be grouping individuals according to a few income ranges. In some sense, this is equivalent to focusing on a typical individual for each income range.

3 Estimation Methods

The estimation problem in the context of the valuation model presented in Section 2 is to use information on the indicator y and the observed covariates x to recover the parameters of the valuation function or the utility difference function.[3] The traditional estimation approach is to assume that the random term ε has distribution function F_ε, so that the probability of a "yes" answer is

$$P(\theta) = \Pr\{z(x,t;\theta) \geq \varepsilon\} = F_\varepsilon(z(x,t;\theta)), \tag{5}$$

[2] See Manski (1986) and Horowitz (1993*a*).

[3] This estimation problem is often referred to as structural discrete choice model. It contrasts with reduced form models where only choice probabilities are estimated. The problem of recovering the structural parameters is treated in the literature under the label of identification. See Manski (1988).

reducing the estimation problem to the maximization of a log-likelihood function with a general form given by

$$\log L = \sum_{i=1}^{N} y_i \log\left[P_i(\theta)\right] + (1-y_i)\log\left[1-P_i(\theta)\right]. \tag{6}$$

In this article, two estimation approaches based on (6) and (5) are considered. The first is the usual parametric approach where F_ε is assumed to belong to some parametric family. The other corresponds to a distribution-free semiparametric approach where the $P_i(\theta)$ is substituted by a nonparametric estimate. Each one of these approaches are discussed in the remainder of this section.

3.1 Censored Logit

Clearly, for the utility difference model, the maximization of (6) leads to standard logit and probit when F_ε is assumed to be logistic or normal, respectively (Hanemann 1984). For the valuation function model the presence of the threshold value t leads to an analog of the censored regression model where the scale of the model can be identified. For logistic F_ε and $s(x,\theta) = x'\beta$, for instance, Cameron (1988) shows that the log-likelihood (6) can be written as

$$\log L = \sum_i (1-y_i)\left[\frac{t_i - x_i'\beta}{\sigma}\right] - \log\left[1+\exp\left(\frac{t_i - x_i'\beta}{\sigma}\right)\right], \tag{7}$$

where σ is a scale parameter.

The expected gains conditional on x, given by equation (4) can be easily computed for logistic F_ε and the estimated β and σ. It is important to note, however, that if F_ε is misspecified and/or the *iid* error assumption is violated, the results obtained in this parametric setting are likely to be poor. In fact, as the results presented in Section 4 suggest, logit estimates may lead to the undue acceptance of a project.

3.2 Klein & Spady Estimator

The basic idea of Klein and Spady's (1993) estimator is to replace $P_i(\theta)$ in (6) with a nonparametric estimate obtained through kernel density estimation. The key development for defining their estimator is to write the $P_i(\theta)$ in terms of estimable densities. Specifically, for any real z, the true probability of a "yes" answer can be written as

$$P(\theta) = \frac{P\,g(z\,|\,y=1)}{P\;g(z\,|\,y=1) + (1-P)\,g(z\,|\,y=0)}. \tag{8}$$

where P is the unconditional probability of $y = 1$, and $g(z\,|\,y)$ is the conditional density of the index z given y. Because $P\,g(z\,|\,y) \equiv g(y, z)$, only estimates of $g_{zy} = g(y, z)$ are needed.

Klein & Spady (1993) propose getting these estimates using the following kernel estimator:

$$\hat{g}_{zy}(z_i; \theta, \hat{\lambda}_y; h_N) = \frac{1}{N-1}\sum_{j\neq i}^{N} \frac{\mathbb{1}(y_j = y)}{h_N \hat{\lambda}_{yj}} K\left[\frac{z_i - z_j}{h_N \hat{\lambda}_{yj}}\right]. \tag{9}$$

The the kernel function $K(\nu)$ is symmetric, integrate to one, have bounded second moment, and must satisfy a some conditions regarding its derivatives. The argument h_N is a nonstochastic sequence of bandwidths satisfying $Nh_N^6 \to \infty$ and $Nh_N^8 \to 0$ as $N \to \infty$. Finally, $\hat{\lambda}_{yj}$ control the bandwidth and define the type of kernel smoothing used. For bias reducing kernels $\lambda = 1$ and the bandwidth is fixed across observations. For locally smoothed kernels λ is a function of a preliminary density estimate, and the bandwidth varies across sample points according to the mass on each of them.[4] For future reference, the estimator defined by the maximization of the quasi-likelihood function obtained by substituting (8) and (9) into (6) will be called Klein and Spady Estimator (KSE).

For technical reasons, Klein & Spady (1993) consider a trimmed version of the estimator. Trimming is necessary to guard against "too" small densities affecting convergence rates. These factors are crucial for the derivation of the asymptotic properties of the estimator. However, as noted by Klein and Spady, the trimming seems to have little effect on estimates. As a result, because of the considerable extra amount of computations required, the trimming factors are often ignored in applied studies (Horowitz 1993*b*). In the application presented in Section 4, it was found that trimming has very

[4] See Klein & Spady (1993) and Silverman (1986, pp. 66-70).

little effect, corroborating this conjecture. For this reason, only the untrimmed version is presented here.

The trimmed estimator is shown to have all the desired properties: consistency, root-N normality and attains the efficiency bound of Cosslett (1987). Monte-Carlo evidence presented by Klein & Spady (1993) indicates that the small sample behavior of the estimator is good, with modest efficiency losses relative to maximum likelihood with known disturbances' distribution. Moreover the KSE is perfectly analogous to the standard maximum likelihood methods. Thus, the information matrix may be taken as the asymptotic covariance matrix and we can also perform likelihood-ratio tests.

Another important feature of the KSE is that it can accommodate heteroscedasticity just by redefining the assumed data generation scheme given in (3) as

$$y = \begin{cases} 1 & \text{if } z(x,t;\theta) \geq h\,(x,t;\theta)\,\varepsilon \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

In this case, provided that h is a known function, bounded away from zero and satisfying some conditions related to model identification and to an index restriction, one can redefine the left-hand side of the inequality in (10) as z/h and proceed just as in the standard specification.

Klein & Spady (1993) have shown that the more general case where h is unknown, but depends on $\{x,t\}$ only through the index $z(x,t;\theta)$ and ε is independent of x and t, can also be accommodated. These are certainly restrictive assumptions, as it limits the forms of unspecified heteroscedasticity in the model. However, there are some instances where it seems to be a reasonable assumption. One such instance is a preference uncertainty context where bid values close to the underlying valuation are assumed to be associated with larger variances due to some sort of ambivalence, as discussed by Ready, Whitehead & Blomquist (1995) among others. For this reason, in this article the KSE results are interpreted as incorporating any sort of unobserved heterogeneity.

To adapt the KSE to valuation models, it is necessary to discuss the model specification and identification. Consider a linear valuation function, with the index given by $\alpha + x_i'\beta - t_i$. As Klein & Spady (1993) have shown, the intercept α is not identified and β is identified only up to a scale parameter. Normalizing the index by multiplying it through by $\kappa \equiv 1/\sigma$, where σ is a scale factor, the probability of a "yes" answer can be written as

$$\Pr(x_i'\gamma - \kappa\, t_i \geq \eta) \equiv F_\eta(x_i'\gamma - \kappa\, t_i) = F_\varepsilon(\alpha + x_i'\beta - t_i) \tag{11}$$

where $\eta = \kappa(\varepsilon + \alpha)$ and $\gamma = \kappa\,\beta$. Clearly, the parameters γ and κ are identified and thus can be readily estimated through the KSE. Once the estimates $\hat{\gamma}$ and $\hat{\kappa}$ are obtained, the valuation function parameters β can be estimated by $\hat{\beta} = -\hat{\gamma}/\hat{\kappa}$. To estimate the valuation function, it remains to determine α which is assimilated into the random error.

The problem of estimating α in this context is analog to the case of Cosslett's (1983) estimator considered by Li (1996). In particular, one can substitute $\hat{\kappa}$ and $\hat{\gamma}$ into equations (8) and (9) to estimate F_η. Then, $E(\eta|\, x)$ can be approximated by numerically integrating the resulting $\hat{F}_\eta$ curve. That is, for a sequence of bid values $t = \{t_1, \ldots, t_M\}$ we have that

$$E(\eta\,|\,x) = \int \eta\, dF_\eta \simeq \sum_{i=1}^{M-1} (x'\gamma - \kappa\, t_i)\Delta\hat{F}_\eta(x'\gamma - \kappa\, t_i), \tag{12}$$

where $\Delta\hat{F}_\eta = \hat{F}_\eta(z_{i+1}) - \hat{F}_\eta(z_i)$. Finally, using the fact that $E(\eta|\, x) = \kappa\,\alpha$, and provided that $E(\varepsilon|\, x) = 0$, the intercept α can be estimated by $\hat{\alpha} = E(\eta|\bar{x})/\kappa$, where $\bar{x}$ is the sample mean of x.

Obviously, the approximation suggested in (12) depends crucially on the bid sequence limits. Ideally, we should have $\hat{F}_\eta \simeq 0$ for t_1 and $\hat{F}_\eta \simeq 1$ for t_M. If t_1 and/or t_M are far way from *observed* bid values, the estimation of F_η through equations (8) and (9) is likely to be poor. Therefore, some truncation of F_η might be required when the bid values in the sample are not large and/or small enough.[5] Nonetheless, it is important to note that the need for some truncation is due to the bid design and not

[5] For more details on the truncation in nonparametric valuation studies see Kriström (1990).

a limitation inherent to the semiparametric method. Estimation of the tails of the distribution is possible as far as there are bid values carrying relevant information about these portions of the distribution.

Given an estimate of asymptotic covariance matrix of γ and κ, the covariance matrix of $\hat{\beta}$ can be easily estimated using the δ-method. For the variance of $\hat{\alpha}$, however, a better alternative seems to be parametric bootstrap procedure: Given the asymptotic normality of the parameter estimates, $\hat{\gamma}$ and $\hat{\kappa}$, generate parameter vectors from a multivariate normal distribution with location and scale given by the estimated parameter vector and the corresponding covariance matrix. Then, for each parameter vector, compute the corresponding intercept estimate. The resulting set of estimates correspond to the empirical distribution, which can be used to compute standard errors and confidence intervals.

Unlike in the logit case, the expected gain conditional on x and c can not be computed directly from (4) in the KSE model. Nonetheless, the integral in (4) can be approximated using a grid of bids. In particular, given equation (11), estimates of the expected gains can be obtained by

$$\hat{G}^{+}(x,c) = \sum_{i=1}^{M-1} (c - t_i)\, \mathbb{1}(c \geq t_i)\ \Delta \hat{F}_{\eta}(x'\hat{\gamma} - \hat{\kappa} t_i), \tag{13}$$

where t_i is an increasing sequence of costs, preferably not far from the observed bid values. Likewise, x and c defining the evaluation point z_i of equation (9) should not be far away from the sample observed values. The expected losses, denoted by $\hat{G}^{-}(x,c)$, can be estimated by substituting $c < t_i$ for the inequality in the indicator function (13).

4 Valuation of the Doce River Basin

4.1 Survey Design and Data Collection

As noted in the Introduction, the project to be valued refers to the management and improvement of the Doce River Basin, according with federal legislation. It is worth to note that this close relation with actual legislation have the effect of reducing considerably the hypothetical nature of the scenario being presented to the interviewees. This scenario included a brief description of of the current condition of the river basin, the role of the administrative agency to be created, the investment plan and its benefits, and the cost share.

The investment plan considered was elaborated taking into account the specific needs of the Doce river basin. In general the descriptions of these investments are very technical, and their relation to concrete benefits are not direct in most cases. For this reason, the scenarios were designed to focus on benefits rather than description of investments.[6] The technicians responsible for the elaboration of the investment plans identified three basic benefits: maintenance/improvement of domestic tap water supply and sewage collection, reduction in pollution levels of the basin's rivers, and improvement of outdoor activities in the areas surrounding the basin's rivers including some parks.

All components of the survey instrument were pretested in preliminary surveys. As a result, the questionnaire underwent several changes before the final format was reached. The major enhancements in this process were in the scenario reliability and its assimilation by interviewees. In this stage a total of 279 interviews were carried. In each of these interviews, an open-ended elicitation question were presented to each respondent. The answers to these open-ended questions were used later as a reference for the bid range choices and for verification of the sample sizes computed before.

Final sample sizes for both applications were determined statistically using the formula proposed by Mitchell & Carson (1989, p.225), with census data on income as a proxy for willingness to pay. The resulting sample sizes were increased by 30%, reflecting the expected proportion of protest bidders, leading to final sample size of 1802 households.[7]

[6] Investments were grouped in classes and had their technical descriptions substituted by a more general characterization, intended to be understandable to an average citizen

[7] It is important to note that the coefficients of variation obtained in the pretest survey are lower

Subjects, the family head in most cases, were interviewed in person by trained personnel from a major Brazilian polling company. In addition to demographic information about the household, the questionnaire collected attitudinal and behavioral information on topics related with the projects, such as pollution, outdoor activities and shortage in water supply. The elicitation question was formulated in the *referendum* format, as described before. The bid value presented to each individual was randomly drawn from a pool of 10 bid values. After the elicitation question, a screening question for those who answered *no* was presented. Those individuals who gave answers like *"I do not believe the money will be used in the projects"* to this screening question were labeled as protest bidders (25.7%) and dropped from the sample.

4.2 Estimation Results

The estimation of the KSE model and the censored logit were implemented in S-PLUS, using the standard normal density as the kernel function and locally smoothed kernels in the former.[8] The same set of covariates were included in both models: the monthly income of the household in thousands of Reais, *income*, the age and years of shooling of the head of the family in years, *age* and *schooling*, respectively. In order to facilitate interpretation, covariates were centered at the sample means at the estimation stage. As a result, the intercept estimates correspond to the estimated willingness to pay conditional on the sample mean of the covariates. Table 1 shows the results obtained.

For each method, Table 1 lists the coefficient estimates and the corresponding standard errors. The first two columns show the results obtained with the censored logit model, while the KSE results are shown in the last two columns. Except for the intercept estimate in the KSE, reported standard errors are estimates based on the expected Hessian evaluated at the coefficient estimates, using the δ-method approximation for the KSE estimates. The standard error of intercept of the KSE model was obtained

than those for income. Therefore, the sample sizes are likely to be larger than required.

[8] The code is available on request. Nonetheless, it is worth to note that Klein and Spady's (1993) estimator is packaged in LIMDEP, facilitating its application by other researchers.

Table 1: Coefficient Estimates

	Logit		KSE	
	Coef	s.e.	Coef	s.e.
Intercept	8.17	0.40	7.87	2.38
Income	5.42	0.86	4.17	0.65
Age	-0.04∗	0.03	-0.02∗	0.02
Schooling	0.28†	0.12	0.14‡	0.08
Scale	6.48	0.07	9.58	2.95

∗ not significant; † significant at the 2% level; ‡ significant at the 7% level. All other coefficients are significant at less than the 1% level.

through the parametric bootstrap procedure described before.

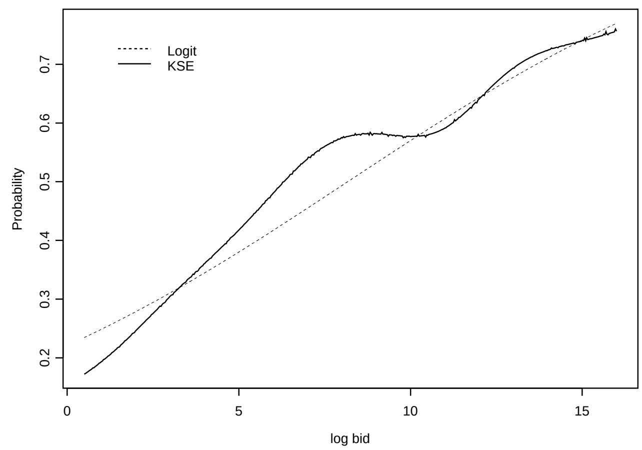
Inspection of Table 1 reveals that the censored logit and the KSE coefficient estimates have the same sign: positive for *income* and *schooling* and negative, but not significant, for *age*. The estimated mean willingness to pay, conditional on the sample mean of the covariates, are very similar according to the logit and the KSE models.[9] The effect of *income* is slightly higher according to the logit model, while the effects of *age* and *schooling* according to the KSE are approximately half the effect in the logit model.

More significant differences are found in the estimated standard errors, which tend to be much larger in the KSE. Nonetheless, the significance of the coefficients is not changed from one method to the other. Obviously, increased variance is the price to be paid for the less strident assumptions of semiparametric methods. However, the increased variance might also be related to heterogeneity structure that can not be captured by the parametric model. This conjecture is corroborated by the shape of the willingness-to-pay distribution implied by each estimation method, as illustrated in Figure 1 with all covariates fixed at the sample means.

The most striking aspect of Figure 1 is the clear indication of bimodality of the willingness-to-pay distribution according to the KSE model. Because there is a signif-

[9] Recall that since covariates were centered at sample means, the estimated willingness to pay conditional on sample means correspond to the intercept estimates.

Figure 1: Willingness-to-pay Function



icant difference in the lower tail, we can conclude that the KSE moves mass from the center of the distribution to the lower tail while the logit insists on placing it symmetrically. This pattern of mass allocation have important consequences for the estimated expected gains and losses. Specifically, moving mass from the center to the lower tail is equivalent to decrease, relative to the logistic distribution, the willingness to pay of some individuals. Thus, one can expect that the logit model will tend to overestimate gains and underestimate losses associated with higher costs, resulting in an overestimate of the net benefits of the project. This overestimation may have important effects on project evaluation, as shown in Table 2.

Table 2 gives estimates of the aggregate gains and losses according to the logit and KSE models. Each column of Table 2 correspond to a hypothetical cost share. All figures, except for 8.17, which corresponds to the overall mean willingness to pay according to the logit model, were arbitrarily chosen. To facilitate a policy oriented discussion, income range specific estimates are provided. For each income range, there are two lines corresponding to the logit (L) and the KSE estimates (K). The bottom lines show unconditional benefit estimates, corresponding to the sum of the benefits over income ranges. Estimates of aggregate gains (losses) correspond to the estimate

Table 2: Aggregate Gains and Losses Estimates

(1000 R$ per month)

Income Range		Individual Cost Share – R$ per month									
R$ per month		3.00		7.00		8.17		10.00		15.00	
Less than 224	K	1333	-198	590	-895	487	-1092	273	-1608	25	-3341
	L	1576	-233	731	-894	566	-1179	369	-1710	99	-3634
224 to 560	K	573	-51	228	-284	190	-357	125	-485	10	-1099
	L	722	-62	356	-252	281	-338	188	-502	54	-1133
560 to 1120	K	408	-12	161	-90	118	-134	81	-197	7	-461
	L	552	-17	304	-77	248	-106	175	-166	57	-423
1120 to 2240	K	397	-2	188	-16	140	-27	83	-55	10	-162
	L	587	-3	381	-16	328	-23	253	-40	105	-127
More than 2240	K	150	0	100	-1	83	-1	59	-1	6	-10
	L	285	0	224	0	206	-1	178	-1	109	-5
Unconditional	K	2860	-263	1267	-1286	1019	-1610	621	-2346	58	-5074
	L	3723	-316	1996	-1239	1628	-1647	1164	-2419	425	-5322

Notes: Gains are presented with positive sign and losses with negative sign. At the time of the survey, R$ 1.15 $\simeq$ US$ 1. L = Censored logit, K = Klein and Spady Estimator.

of individual gains (losses) times the number of gainers (losers) in each income range.[10] Logit estimates of individual gains and losses were computed directly from equation (4), with F_ε defined by the coefficient estimates $\hat{\beta}$ and $\hat{\sigma}$ presented in Table 1. KSE individual estimates were obtained from equation (13), using the estimates $\hat{\gamma}$ and $\hat{\kappa}$ presented in Table 1 and grid t_i with 1000 points equally spaced between 0.5 and 15. In both cases, the vectors of covariates were fixed at the sample means within each income range.

The results given in Table 2 support the claim that the logit model tend to overestimate the net benefits in this application. Interestingly, the extent of the overestimation increases with income. For the lower income ranges, logit net benefit estimates are relatively close to the KSE estimates. As we move to higher ranges, we observe that the difference between them tend to increase. The overall consequence of these findings is

[10] The number of gainers and losers were obtained by multiplying the number of households in the population by the proportions implied by the logistic distribution and equation (8). In order to facilitate using census data, income ranges for the counting of households were defined according to the income of the head of the family, instead of the household income.

Table 3: Example of Project Financing

Income Range	Tax Price	Gains	Losses	Net Benefit
Less than 224	3.00	1333	198	1135
224 to 560	3.00	573	51	522
560 to 1120	7.00	161	90	71
1120 to 2240	10.00	83	55	28
More than 2240	15.00	6	10	-4
Total		2156	404	1752

Notes: At the time of the survey, R\$ 1.15 $\simeq$ US\$ 1.00.
All figures in R\$ per month.

that, even though the logit tends to underestimate the burden for both the richer and poorer, it does more heavily for the former.

Certainly, the most important result presented in Table 2 is the apparent leniency of the logit (relative to the KSE estimates) regarding the project evaluation. As expected, the unconditional estimates indicate that charging the mean willingness to pay estimate (R\$ 8.17) produce net benefit close to zero according to the logit.[11] Thus, any project with average cost smaller than R\$ 8.17 would pass the Kaldor-Hicks criterion. However, according to the KSE estimates, the net benefit at this cost is significantly negative. In fact, it is close to zero only at a cost of R\$ 7.00 per month. Thus, any project with average cost between 7.00 and 8.17 Reais would be unduly accepted were the logit estimates used in the project analysis. This finding illustrates the importance of allowing a more general distribution of benefits when formulating financing policies.

To conclude this section, it is interesting to evaluate the project considering alternative financing plans. The case of a flat tax price analyzed above is certainly very helpful for the project analysis. However, it is often the case that policy makers are interested in progressive tax prices. Table 3 gives a hypothetical financing scheme for the Doce River Basin project, with tax prices differentiated by income ranges. The first column shows the tax price for each income range. For instance, there is a monthly charge of R\$ 3.00 for households with income less than R\$ 336, of R\$ 5.00 for households with income between R\$ 336 and R\$ 760, and so on. The second and third columns show

[11] The small difference observed is due to approximation error.

the aggregate gains and losses, according to Table 2. The fourth column shows the aggregate net benefit. Thus, any project with a total cost smaller than R$ 1.8 million per month would be justifiable given this financing scheme.

5 Conclusion

This article considered the valuation of a project for the improvement of water resources in Brazil and proposed the application of Klein and Spady's (1993) semiparametric estimator and related methods to contingent valuation models. Results obtained indicate that the usual censored logit approach produces good estimates of the conditional mean willingness to pay, but it fails to capture a rich heterogeneity structure. Specifically, the proposed semiparametric approach suggests that the willingness-to-pay distribution is bimodal, while the logit insists to place mass symmetrically about the mean.

In this application, the bimodality on the welfare analysis of the project has an important effect when the Kaldor-Hicks criterion is used. Even though the estimates of the overall conditional mean willingness to pay are similar, the logit insistency in allocating mass symmetrically lead to a significant overestimation (relative to the semiparametric method) of net benefits. As a result of this overestimation, the logit might lead to the undue acceptance of projects.

Even though the results obtained can not be generalized, the evidence of bimodality suggests that the usual logit approach can be usefully complemented with semiparametric methods. If the shape of the distribution implied by the semiparametric model is in line with the logistic assumption, confidence about the results obtained through parametric methods is strengthened. However, if the semiparametric model suggests severe deviation from the logistic distribution, parametric methods should be viewed with care, specially when the main interest is not the overall conditional mean benefits. In such a case, the semiparametric approach seems to provide a more accurate representation of the heterogeneity structure, enriching welfare analysis.

References

Cameron, T. A. (1988), 'A new paradigm for valuing non-market goods using referendum data: Maximum likehood estimation by censored logistic regression', *Journal of Environmental Economics and Management* **15**(3), 355–380.

Cameron, T. A. & James, M. D. (1987), 'Efficient estimation methods for "closed-ended" contingent valuation surveys', *Review of Economics and Statistics* **69**(2), 269–276.

Chen, H. Z. & Randall, A. (1997), 'Semi-nonparametric estimation of binary response models with an application to natural resource valuation', *Journal of Econometrics* **76**(1-2), 323–340.

Cosslett, S. R. (1983), 'Distribution-free maximum likelihood estimator of the binary choice model', *Econometrica* **51**(3), 765–782.

Cosslett, S. R. (1987), 'Efficiency bounds for distribution-free estimators of the binary choice and censored regression models', *Econometrica* **55**, 765–782.

Creel, M. & Loomis, J. (1997), 'Semi-nonparametric distribution-free dichotomous choice contingent valuation', *Journal of Environmental Economics and Management* **32**, 341–358.

Hanemann, W. M. (1984), 'Welfare evaluations in contingent valuation experiments with discrete responses', *American Journal of Agricultural Economics* **66**(3), 332–341.

Horowitz, J. L. (1992), 'A smoothed maximum score estimator for the binary response model', *Econometrica* **60**(3), 505–531.

Horowitz, J. L. (1993*a*), Semiparametric and nonparametric estimation of quantal response models, *in* G. S. Maddala, C. R. Rao & H. D. Vinod, eds, 'Handbook of Econometrics', Vol. 11, Elsevier Science Publishers, chapter 2, pp. 45–84.

Horowitz, J. L. (1993*b*), 'Semiparametric estimation of a work-trip choice model', *Journal of Econometrics* **58**(1-2), 49–70.

Klein, R. W. & Spady, R. H. (1993), 'An efficient semiparametric estimator for binary response models', *Econometrica* **61**(2), 387–421.

Kriström, B. (1990), 'A non-parametric approach to the estimation of welfare measures in discrete response valuation studies', *Land Economics* **66**, 135–139.

Li, C. (1996), 'Semiparametric estimation of the binary choice model for contingent valuation', *Land Economics* **72**(4), 462–73.

Manski, C. F. (1975), 'Maximum score estimation of the stochastic utility model of choice', *Journal of Econometrics* **3**(3), 205–228.

Manski, C. F. (1986), 'Semiparametric analysis of binary response from sample-based samples', *Journal of Econometrics* **31**, 31–40.

Manski, C. F. (1988), 'Identification of binary response models', *Journal of the American Statistical Association* **83**(403), 729–738.

McConnell, K. E. (1990), 'Models for referendum data: The structure of discrete choice models for contingent valuation', *Journal of Environmental Economics and Management* **18**(1), 19–34.

Mitchell, R. C. & Carson, R. T. (1989), *Using Surveys to Value Public Goods: The Contingent Valutation Method*, Resources for the Future, Washington, D.C.

Ready, R. C., Whitehead, J. C. & Blomquist, G. C. (1995), 'Contingent valuation when respondents are ambivalent', *Journal of Environmental Economics and Management* **29**(2), 181–196.

Silverman, B. W. (1986), *Density Estimation for Estatistics and Data Analysis*, number 26 *in* 'Monographs on Statistics and Applied Probability', Chapman & Hall/CRC, New York.