Whitehead, J.C., Groothuis, P.A., and Blomquist, G.C. (1993) Testing for Nonresponse and Sample Selection Bias in Contingent Valuation: Analysis of a Combination Phone/Mail Survey, **Economics Letters**, 41(2): 215-230. Published by Elsevier (ISSN: 0165-1765).

Testing for non-response and sample selection bias in contingent valuation: Analysis of a combination phone/mail survey

John C. Whitehead, Peter A. Groothuis, and Glenn C. Blomquist

ABSTRACT

We use a combination phone/mail survey to test for possible sample biases in contingent valuation. We find no sample selection bias hut do find non-response bias. We show how failure to correct for non-response bias distorts aggregate benefit estimates.

1. INTRODUCTION

Contingent valuation (CV) mail surveys are used to collect primary data for estimating the value of environmental resources. One problem with mail surveys is non-response. Response rates on CV mail surveys typically range between 20% and 00%. Drawing inference about the population from a sample with non-response may generate biased results.

Non-response bias arises if non-respondents differ from respondents in observable characteristics that influence willingness to pay (WTP). Non-response can lead to sample selection bias even if non-respondents arc similar to respondents in observable characteristics but differ in their WTP for environmental preservation due to unobservable characteristics.

Mitchell and Carson (1989, p. 277) suggest that individuals who feel strongly about an environmental amenity will be more likely to respond to a survey. This would be an example of sample selection bias that biases estimated WTP upward. If CV mail surveys suffer from either ordinary non-response bias or sample selection bias, the generalization of individual WTP values to the population will produce biased aggregate benefit estimates.

Non-response and sample selection bias are well-known problems among CV researchers. Tests for bias have been scarce, however, because data on non-respondents, which is necessary to conduct the tests, has not been available. Two studies have addressed the bias problem, although the authors were hindered with less than ideal data for the necessary tests [Edwards and Anderson (1987), Loomis (1987)].

In this paper, we use data on non-respondents obtained by employing a two-stage phone sampling and mail survey procedure. This procedure is superior to a one-stage mail survey because information on non-respondents is obtained. Using this data, we test for non-response and sample selection bias and illustrate the effect of bias on aggregate benefit estimation.
We proceed on the basis of Mitchell and Carson's (1989) assessment that if a survey is designed and implemented well, the CV method can produce valid and reliable estimates of the value of environmental commodities. While this is the prevailing view, we recognize that the validity of the CV method is the subject of a vigorous debate, see Kahneman and Knetsch (1992) and Smith (1992), for example.

2. SAMPLE DESIGN AND SURVEY

A CV survey was designed to measure the value of wetlands preservation in the western Kentucky coal field [Blomquist and Whitehead (1990)]. An important feature of the survey was a combination phone interview/mail questionnaire which gathered information on almost every sampled household. In the phone interview we collected socioeconomic data. Phone respondents were asked if they would complete a mail questionnaire concerning wetlands. If yes, they gave their names and addresses for the CV mail survey list. Of the 926 people called, 730 (79%) completed the phone interview and 641 (69%) gave their names and addresses for the mail survey. Questionnaires were sent to all 641 households who gave their names and addresses. Mail survey procedures followed Dillman (I978) with a primary instrument mailing, one postcard reminder and two replacement survey instruments. The total number of replies was 487 - 67% of the households who participated in the phone survey and 76% of those households mailed a questionnaire.

Given this survey method, non-respondents are of three types: unit non-response, item non-response, and protest non-response. Unit non-response occurs at two levels. At the phone interview level, unit non-response occurs when no information is obtained on the unit of analysis. Unit non-response also occurs at the mail level if individuals decline to submit to the mail survey or decline to return the questionnaire. Item non-response occurs when the individual fails to respond to the contingent market value elicitation question. Protest non-response to the value elicitation question, determined by follow-up questions, are selected out of samples by CV researchers [Mitchell and Carson (1989)].

Table 1
Variable descriptions

Variable names	Description
Tenure	Number of years living at current address
Wetland knowledge	Respondent had prior knowledge of wetlands in Kentucky (Knowledge = 1, No knowledge = 0)
Age	Respondent's age
Education	Years of education of the respondent
Children	Number of children in household
Gender	Gender of respondent (Male = 1, Female = 0)
Information	Information is an index measuring the number of wetland characteristics presented in the survey instrument
Conservationist	Respondent is a member or donates money to conservation organizations (Member = 1, Not a member = 0)
Income	Annual household income in thousands (1990 dollars)
Polychotomous choice	Dummy variable that controls for the number of choices in the contingent market (*Polychotomous choice* = 1, *Dichotomous choice* = 0)

3. TESTING FOR SAMPLE NON-RESPONSE BIAS

Non-response bias can be tested by comparing characteristics of respondents who returned completed surveys and non-respondents who failed to return a completed survey. Table 1 contains descriptions of variables used in the study. Table 2 shows variables which revealed significant differences between mail respondents and mail non-respondents using a difference in means test. Respondents have more wetland knowledge, are younger, have more education, and are more likely to be male than non-respondents to the mail survey.

These statistically significant differences in means indicate that if these variables influence WTP values, then WTP estimates will be biased. To correct the problem of non-response bias, aggregate WTP measures should use the population means when available and not the sample means.

Table 2
Significant differences in socioeconomic characteristics of respondents vs. mail non-respondents

Variable	Respondents	Non-respondents
Wetland knowledge	72.30%	58.19%
	(44.81)	(49.43)
Age	48.17	52.06
	(17.38)	(21.59)
Education	12.87	10.88
	(2.60)	(3.03)
Gender	48.20%	34.91%
	(50.00)	(47.78)
Sample size	361	232

Pairs of variables shown are significantly different at, at least, the 0.05 level.
Standard deviation in parentheses.
The differences in means of other variables were insignificant at the 0.05 level. The means are tenure (14.92, 15.03), children (0.68, 0.68), income (–, 25.92), conservationist (–, 0.20), log of specified amount (2.46), polychotomous choice (–, 0.44), and information (–, 12.32). The first mean listed is the mail non-respondents, the second mean is the mail respondents. Blanks indicate means were obtained through mail questionnaire and not available for non-respondents.

4. TESTING FOR SAMPLE SELECTION BIAS

To test for sample selection bias in our data, we employ the bivariate probit technique [Dubin and Rivers (1989)]. In our model the selection equation's dependent variable, *USABLE,* is equal to one if an individual returned a completed survey and zero otherwise. Independent variables include the respondent's gender, age, tenure, education, the number of children and wetland knowledge.

In the WTP equation, the dependent variable, *YES,* is equal to one if the individual is willing to pay the specified amount and zero if unwilling to pay that amount. The specified amount ranged randomly from $3 to $49. The independent variables in the WTP model include the log of the specified amount to be paid (price), gender, age, education, income, the number of children, tenure, wetland knowledge, information about wetland characteristics, a dummy variable on whether the respondent is a conservationist, and a dummy variable on the number of choices in the contingent market.

Missing data was handled with a data imputation method so that information from the value elicitation question, which may reveal sample selection, is not lost [Little and Rubin (1989)]. Thus respondents who answered the contingent valuation question but failed to respond to other questions were kept by using the mean of the phone survey data, which is the most representative sample. The income non-responses were replaced with values obtained from a (wage equation) regression imputation. Accordingly, our sample size increases from 361 respondents without imputation to 402 with imputation.

Consider a univariate probit model with hypothesized sample selection, where the sample is selected using a univariate probit selection rule. Estimation of such a model is straightforward when the error terms of the two equations are uncorrelated. Univariate pro bit is the appropriate technique. The problem arises when the error terms between equations are correlated because of sample selection bias. In this case bivariate probit with partial observability corrects for sample selection bias.

Univariate and bivariate probit results are presented in Table 3. Comparing the univariate results with the bivariate results shows that no coefficients change in sign; however, two change in significance. Education and tenure are both significant in the univariate pro bit but insignificant in the bivariate.

Table 3
Determinants of the probability of WTP and determinants of probability of response (absolute value of t-statistics in parentheses)

	Univariate		Bivariate	
	USEABLE[a]	*YES*[b]	*USEABLE*[a]	*YES*[b]
Constant	−1.029[e]	−1.233[e]	−1.028[e]	−1.233
	(3.39)	(2.76)	(3.46)	(0.322)
Log amount[c]	–	−0.369[e]	–	−0.369[e]
		(5.04)		(4.76)
Gender	0.119	0.116	0.119	0.116
	(1.19)	(0.83)	(1.18)	(0.52)
Age	−0.008[d]	−0.003	−0.008[d]	−0.003
	(1.78)	(0.57)	(1.79)	(0.29)
Children	−0.060	0.097	−0.060	0.097
	(1.18)	(1.38)	(1.20)	(0.85)
Education	0.101[e]	0.647[e]	0.101[e]	0.646
	(5.86)	(2.73)	(5.83)	(0.40)
Income	–	0.009[e]	–	0.009[e]
		(2.75)		(2.64)
Tenure	0.005	−0.010[d]	0.005	−0.010
	(1.35)	(1.79)	(1.34)	(1.09)
Information	–	0.060[e]	–	0.060[e]
		(2.75)		(2.64)
Wetland knowledge	0.243[e]	0.020	0.243[e]	0.020
	(2.30)	(0.13)	(2.26)	(0.05)
Conservationist	–	0.641[e]	–	0.641[e]
		(3.58)		(3.67)
Polychotomous choice	–	0.457[e]	–	0.457[e]
		(3.27)		(3.09)
Rho	0		−0.290	
			(0.01)	
Log likelihood	−472.44	−231.57	−704.00	
Sample size	730	402	730	402

[a] Dependent variable is equal to one if the individual returned a completed survey.
[b] Dependent variable is equal to one if the individual is willing to pay the specified amount.
[c] *Log amount* is the amount of money individuals were asked to pay in the dichotomous choice question.
[d] Significant at the 0.10 level.
[e] Significant at the 0.05 level.

In the selection equation, *USABLE,* we find that age lowers the likelihood of response, while education and wetland knowledge increase the likelihood of response. In the WTP equation, *YES,* the log of the specified amount to be paid has a negative effect on the probability of being willing to pay while income, information, and being a conservationist have a positive influence.

The rho statistic reported in Table 3 is the correlation between error terms in the selection and response equation. It is constrained at zero when univariate probit is used. Selection bias is then tested for using the Wald and the likelihood ratio test [Dubin and Rivers (1989)J. Both tests find no sample selection bias. Rho is not significantly different than zero. Univariate probit analysis is therefore acceptable for calculating WTP estimates.

5. SAMPLE BIAS AND AGGREGATE BENEFIT ESTIMATION

Often the primary purpose of a CY study is to obtain individual and aggregate WTP estimates. Willingness to pay is estimated according to the method of Cameron and James (1987) for dichotomous choice data and the probit technique. Using the full sample means, the point estimate of household WTP is $6.00 to preserve a wetland area. Using the mail sample means, the point estimate of household WTP is $8.01. Non-response biases WTP upwards by 33%. This bias is due to differences in observable characteristics and not sample selection bias due to unobservable characteristics.

To illustrate the magnitude of the bias across the relevant population we aggregate the full sample and mail sample household WTP estimates for Kentucky. With the 1990 Kentucky population of about 1,400,000 households, aggregate WTP ranges from $11.2 million using the mail sample means to $8.4 million using full sample means. Not correcting for non-response bias results in upward bias in WTP of $2.R million. Biased WTP estimates used to calculate aggregate WTP could lead to mistakes in benefit-cost analysis and inappropriate policy analysis.

6. IMPLICATIONS FOR FUTURE CV RESEARCH

Non-response and sample selection bias can both be problems in CV mail surveys of general populations. With data on both respondents and non-respondents to a combination phone/mail CV survey about Kentucky wetlands we are able to test for both ordinary non-response bias and sample selection bias. In our study, we find non-response bias on observable characteristics such as education. We fine no sample selection bias on unobservable characteristics. Our WTP estimates are then corrected for non-response bias. Failure to correct for non-response bias would distort our aggregate benefits of wetland preservation upwards by 33%. If possible, future CV surveys should obtain data on non-respondents, test for both non-response and sample selection bias, and correct for both biases if found. The combination phone/mail design is a useful survey method which facilitates such corrections.

REFERENCES

Blomquist, G.C. and J.C. Whitehead, 1990, The effect of alternative and reclaimed areas on the value of wetlands, Final Report (U.S. Department of the Interior. Office of Surface Mining Reclamation and Enforcement).

Cameron, T.A. and M.D. James, 1987, Efficient estimation methods for 'closed-ended' contingent valuation surveys. The Review of Economics and Statistics 69, 269-276.

Dillman. D.A .. 1978, Mail and telephone surveys: The total design method (John Wiley. New York. NY).

Dubin, J.A. and D. Rivers, 19R9. Selection bias in linear regression. logit and pro bit models. Sociological Methods and Research 18, 360-390.

Edwards, S.F. and G.D. Anderson. 1987, Overlooked biases in contingent valuation surveys: Some considerations, Land Economics 63, 16S-17S.

Greene. W. 1989. LIMDEP: User's manual (Econometric Software, Inc.).

Kahneman, D. and J. Knetsch, 1992, Valuing public goods: The purchase of moral satisfaction, Journal of Environmental Economics and Management 22, 57-70.

Little, R.J.A. and D. Rubin, 1989, The analysis of social science data with missing values, Sociological Methods and Research IS. 292-323.

Loomis. J.B .. 1987, Expanding contingent value sample estimates to aggregate benefit estimates: Current practices and proposed solutions, Land Economics 63. 396-402.

Mitchell, R.C. and R.T. Carson, 19R9. Using surveys to value public goods: The contingent valuation method (Resources for the Future. Washington, DC).

Smith, V.K., lY92. Arbitrary values, good causes, and premature verdicts, Journal of Environmental Economics and Management 22, no. l. 57 70.