Environmental and Resource Economics **11:** 107–116, 1998.

Construct Validity of Dichotomous and Polychotomous Choice Contingent Valuation Questions

JOHN C. WHITEHEAD[1], GLENN C. BLOMQUIST[2], RICHARD C. READY[3] and JU-CHIN HUANG[1]
[1]*Department of Economics, East Carolina University, Greenville, North Carolina 27858, USA;* [2]*Department of Economics and Martin School of Public Policy and Administration, University of Kentucky, Lexington, KY 40506, USA;* [3]*Department of Economics and Social Studies, Agricultural University of Norway, N-1432 Ås, Norway*

Accepted 4 April 1997

Abstract. In this note we conduct construct validity tests for dichotomous choice (DC) and polychotomous choice (PC) contingent valuation questions. Contrary to previous results, we find that DC and PC estimates of willingness to pay are theoretically valid, convergent valid, and similar in terms of statistical precision. Similar to previous results, PC respondents are less sensitive to information than DC respondents. We conclude that DC and PC valuation questions are construct valid for this study. Sequential PC valuation questions could be used in studies where obtaining information about the certainty or intensity of respondent preferences would be useful.

Key words: contingent valuation, dichotomous choice, polychotomous choice, construct validity

1. Introduction

One goal of contingent valuation (CV) research is to obtain construct valid estimates of willingness to pay (WTP) for environmental resource allocation changes. Construct validity includes convergent and theoretical validity. Convergent validity is the extent to which two measures of WTP, measured differently, are correlated. Convergent validity tests can be conducted with WTP estimates measured using revealed behavior and behavioral intentions methodologies (Cameron 1992; Jakus 1994) or different WTP elicitation methods (Boyle and Bishop 1988; Kriström 1993). Theoretical validity is the extent to which a measure behaves according to theoretical predictions. One type of theoretical validity test employs split-samples to determine if respondents consider contingent market information in theoretically predicted ways (Boyle et al. 1994; Carson and Mitchell 1995). Another type of theoretical validity test focuses on the internal consistency of CV responses (Whitehead et al. 1995).

Research concerning WTP elicitation methods has generally led CV researchers to favor dichotomous choice (DC) valuation questions (Mitchell and Carson 1989; Arrow et al. 1993; Freeman 1993). DC questions are relatively similar to actual

market transactions which increase respondent familiarity with the contingent market and validity of WTP. Also, since they require only a yes or no response, DC questions tend to be relatively easy to answer. With polychotomous choice (PC) questions respondents are given a multiple choice. One rationale for asking PC questions is that they provide more information to the researcher than do DC questions. In addition to knowing whether the respondent would be willing to pay or not we learn something about the strength or certainty of the underlying preference. Also, the NOAA Panel on Contingent Valuation (Arrow et al. 1993) has recommended that DC WTP formats include a middle response, a form of PC question, in order to enhance the reliability of CV for estimation of passive use values.

PC questions are less similar to actual market transactions, relative to DC questions, and differences in response patterns may emerge. In order to confidently use PC data for policy purposes, the resulting WTP estimates should be construct valid. Willingness to pay should be similar when estimated with DC or PC data, or differ for explainable reasons, and WTP with PC data should be explained by covariates in predicted ways. In the initial test of the validity of PC data, Ready et al. (1995) presented DC valuation questions and six choice PC valuation questions to split-samples in two mail surveys. They find that in both data sets DC and PC WTP estimates are not convergent valid; PC WTP estimates exceed DC WTP estimates. Also, PC WTP estimates were less theoretically valid and measured with less statistical precision than DC WTP estimates. In similar research, Johannesson et al. (1993), Li and Mattsson (1995) and Swallow et al. (1993) have investigated the econometric properties of PC data but not their validity.[1]

In this note we test the construct validity of DC and PC contingent valuation questions which focus on improving the environmental quality of the Pamlico Sound in North Carolina. The Pamlico Sound is part of the Albemarle-Pamlico estuarine system which is the second largest estuarine system in the United States. We use a different questioning mode, survey administration, and number of potential middle responses than has been used before with PC questions. We first describe these differences. Next we compare the theoretical validity of DC and PC data through internal consistency tests, we test for convergent validity by comparing benefit estimates from the different value elicitation methods, and we compare the statistical precision of WTP estimates. We then compare the effects of an information treatment about quality and related goods on DC and PC WTP. Conclusions and suggestions for future research are offered at the end of the paper.

2. The Pamlico Sound Contingent Valuation Study

Survey respondents were presented a contingent market for quality improvements in the Pamlico Sound during a telephone interview.[2] DC and PC willingness to pay questions were presented to split samples. Respondents were first told that agricultural and commercial fishing practices pollute water and destroy wildlife habitat in the Pamlico Sound and that the result during the past ten years has been

decreasing fish catches, disease in crabs, closed shellfish beds, and the disappearance of underwater grasses. Tougher laws that would require farmers to control pollution and restrict some commercial fishing practices were next proposed. The proposed laws were described with a goal of restoring water quality and fish and wildlife habitat.

Respondents were told that tougher laws would lead to higher consumer prices and that enforcement of these laws would lead to higher taxes. The willingness to pay question was closed-ended: "If you knew the money would be used to restore water quality and fish and wildlife habitat in the Pamlico Sound, would you and your household be willing to pay $PT each year, in higher consumer prices and state taxes?" One of six randomly assigned annual price and tax amounts ($PT = 5, 25, 50, 100, 200, or 300) was presented to each respondent. Follow up questions were used to identify protest and outlier responses which were then deleted.

One-half of the respondents received a DC valuation question and the other half received a PC valuation question. The DC treatment was implemented in the usual way; the closed-ended question was presented by the interviewer implying to the respondent that a "yes" or "no" answer was expected. The PC treatment contained four response levels: "definitely yes," "probably yes," "probably no," and "definitely no." The PC response levels were not initially offered to respondents by the interviewer implying that a "yes" or "no" response was expected. If the respondent immediately answered yes or no they were considered to be certain about their answer and it was coded as "definitely yes" or "definitely no." If the respondent hesitated or expressed difficulty with the PC question the interviewer offered the four response categories and the respondent could choose any of the four.

The sequential PC question allows interaction between respondent and interviewer which is not possible with mailed questionnaires and, *a priori*, was expected to reduce the number of uncertain, middle responses chosen, relative to Ready et al. (1995). One reason not to offer middle responses unless the respondent is hesitant is to avoid low cost, noncommittal answers. Ready et al. (1995) argue that respondents who find the middle responses attractive do not incur the costs of formulating a coherent WTP response. By waiting to offer the middle response we might avoid this problem (see Bishop 1987). Thus, the goal of the sequential PC approach is to elicit high quality WTP responses from those who are able to express them but not to force those who can not to overstate their confidence in the response.

Certain types of information in contingent markets may improve the theoretical validity of responses to DC and PC valuation questions (Blomquist and Whitehead, forthcoming). For instance, additional information may cause respondents to give an answer after further considering market characteristics such as resource quality and the number and availability of related goods. Ready et al. (1995) find that PC respondents are less sensitive to contingent market information than DC respondents. Therefore, in the current study one-half of the respondents received additional information preceding the contingent market. These respondents were

told that: "The Pamlico Sound provides habitat for fish and wildlife such as flounder, bay scallops, blue crabs, loggerhead sea turtles, mallard ducks, and brown pelicans. The Albemarle Sound in North Carolina and the Chesapeake Bay in Virginia provide the same kind of fish and wildlife habitat." The first sentence describes additional quality characteristics about the Pamlico Sound relating to use values (seafood, fishing, hunting, wildlife observation) and passive use values (i.e., the existence of loggerhead sea turtles). The second sentence describes two related environmental resources with similar characteristics which may be substitutes or complements to the Pamlico Sound.

The data are from a CV market integrated into a 1993 telephone survey of eastern North Carolina conducted by a university survey research laboratory. The sample was drawn through random digit dialing. Once contact was established with a member of a household, a random selection procedure was used to help maintain a representative distribution of ages and sexes. Of the households that were contacted 1021 completed the questionnaire for a 71% response rate. We employ complete case analysis by dropping all units with item nonresponse on any question leaving 603 cases for a useable response rate of 41%.[3]

3. Empirical Model and Results

The theoretical construct of willingness to pay for a quality improvement is the same whether elicited with DC or PC valuation questions: WTP = $f(p,q^{\circ},q',y; z)$, where p is the own-price of on-site use of the Pamlico Sound, q is the quality of the Pamlico Sound, $q' > q^{\circ}$, y is income, and z represents survey treatments (PC and additional information). For the theoretical construct of WTP to be internally valid, certain theoretical relationships among variables are expected to hold (Whitehead 1995). The effect of income on WTP is positive if quality is a normal good and the effect of the own-price on WTP is negative if quality and recreation trips are complements.[4] For convergent validity the PC treatment should have no effect on WTP and the relationships among variables, including the information treatment, should be the same for DC and PC data. The influence that additional information about resource quality and related goods should have on WTP is unclear (Blomquist and Whitehead, forthcoming).[5]

The validity tests are conducted through binary logistic regression with the PC data treated as dichotomous: $\pi=\{1+\exp[-(\tau PT + x'\beta)]\}^{-1}$, where π is the probability of a yes response, x is a vector of independent variables (p, y, z) and ß is a vector of coefficients. The DC and PC WTP estimates are found by the method of Cameron (1988), dividing $x'\beta$ by the negative of τ, with other independent variables evaluated at their mean. WTP confidence intervals are found using the analytic technique of Cameron (1991) which has been found to be more reliable for small samples when compared to simulation approaches (Cooper 1994).[6,7] The statistical precision of WTP estimates is assessed with the coefficient of variation

which measures, in percentage terms, the relative size of the dispersion of WTP around the mean WTP relative to the mean.[8]

Demographic variables for the complete cases are representative of the population of eastern North Carolina (Table I). Sixteen percent of the respondents had participated in outdoor recreational activities during the twelve months prior to the survey, 94% were at least "somewhat concerned" about Pamlico Sound resources, 83% felt that the laws would be at least "somewhat effective" in achieving these goals, and 49% knew something about the problems, resources, and uses of the Pamlico Sound before the telephone survey.

Final sample sizes and response rates are similar for the PC and DC sub-samples. There are no significant differences between the DC and PC sub-samples for the own-price,[9] income, and additional price and tax ($PT) variables. In addition, there are no statistical differences between the sub-samples for dummy variables measuring whether the respondent had participated in outdoor recreation trips on or around the Pamlico Sound, knowledge about the Pamlico Sound before the survey, concern for Pamlico Sound resources or perceived effectiveness of the program.

To compare the PC and DC responses, the definitely yes (no) and probably yes (no) PC responses were combined to form binary yes and no categories as in Ready et al. (1995). The overall percentage of yes responses is equal between the sub-samples (Table I). This is tested by merging the PC and DC data and including a dummy variable for the PC treatment in the logistic regression model (Table II). We find no statistically significant effect of the PC format on responses to the CV questions.[10]

The second and third models in Table II are used to test for theoretical validity through internal consistency and convergent validity through the effects of the PC treatment on coefficient estimates. The PC responses are theoretically valid according to the own-price and income coefficients while DC responses are theoretically valid according to the income coefficient. The DC and PC models are specified with a dummy variable for the information treatment to test for information effects (Table II). The additional information has a significant positive effect on the proportion of DC yes responses but no significant effect on the PC responses.

While there are differences in individual coefficient estimates, the likelihood ratio test for differences in the vectors of coefficients for the DC and PC data suggests that the two elicitation approaches do not yield different results (χ^2=7.83[5]) and WTP estimated with the DC and PC samples are not significantly different (t=0.80).[11] The DC WTP estimate is more statistically precise since the coefficient of variation is 40% less than that of the PC WTP estimate. The differences in coefficients of variation are suggestive, but not conclusive, since we do not have a formal test of whether such differences are statistically significant.

Table I. Data Summary.

Variable	Pooled	DC	PC
Percent YES	0.64	0.64	0.64[b]
	(0.48)[a]	(0.48)	(0.48)
Dollar Amount (PT)	116.30	110.23	122.47
	(109.89)	(102.66)	(104.78)
Own-Price	112.67	111.77	113.59
	(80.70)	(79.95)	(81.58)
Income	29,714	29,654	29,774
	(20,258)	(19,810)	(20,738)
Information (=1)	0.50	0.50	0.51
	(0.50)	(0.50)	(0.50)
Polychotomous Choice (=1)	0.50	—	—
	(0.50)		
Recreation Participation (=1)	0.16	0.17	0.15
	(0.36)	(0.37)	(0.35)
Concern About Pamlico Sound (=1)	0.94	0.94	0.93
	(0.24)	(0.23)	(0.25)
Effectiveness of Program (=1)	0.83	0.84	0.82
	(0.38)	(0.37)	(0.39)
Respondent Knowledge (=1)	0.49	0.51	0.47
	(0.50)	(0.50)	(0.50)
Sample Size	603	304	299

[a] Standard deviation in parentheses.
[b] The PC data contain 35% "definitely yes," 29% "probably yes," 11% "probably no," and 25% "definitely no" responses.

4. Conclusions

In this paper we conduct construct validity tests for DC and PC CV data. We find that WTP is theoretically valid with both types of data, the estimates are convergent valid, and are similar in terms of statistical precision. PC respondents are less sensitive to information than DC respondents. Overall, we conclude that DC and PC valuation questions are construct valid for this study. The PC valuation question is a potentially valuable technique for eliciting WTP, especially when the intensity or certainty of respondent preferences is an issue that must be considered.

That these results are somewhat different from Ready et al. (1995) is not surprising since we introduced two differences in the PC survey method. In this study we (1) reduced the number of middle response levels offered and (2) altered the method in which the middle responses were offered through a telephone interview.

Table II. Theoretical and Convergent Validity Results[a].

Variable	Pooled	DC	PC
Intercept	0.48**	0.11	0.85***
	(2.11)[b]	(0.37)	(2.82)
PT	–0.0045***	–0.0056***	–0.0035***
	(5.32)	(4.45)	(2.98)
Own-Price	–0.0036*	–0.0022	–0.0056**
	(1.91)	(0.75)	(2.17)
Income	0.000034***	0.000040***	0.000033***
	(4.30)	(3.23)	(3.05)
Information	0.10	0.52**	0.30
	(0.59)	(1.99)	(1.21)
PC(=1)	0.034		
	(0.19)		
Sample Size	603	304	299
χ^2	57.90***	47.30***	20.82***
WTP[c]	$256.45	$234.38	$293.41
	(204.53, 308.37)	(177.05, 291.71)	(186.80, 400.02)
C.V.[d]	12.31%	14.87%	22.09%

*, **, *** indicate that the coefficient is significantly different from zero at the α = 0.10, 0.05, and 0.01 levels.
[a] Dependent variable is $\log[\pi/(1-\pi)]$.
[b] Absolute value of the t-statistic in parentheses.
[c] 90% confidence interval in parentheses.
[d] Coefficient of variation.

By reducing the number and not explicitly offering the middle responses initially, we reduce the incentive to give a careless answer. Both deviations in survey method were designed to increase the thought and care given to their valuations by respondents. Respondents in this study who would have answered "maybe yes" or "maybe no" if these two additional middle responses had been offered, instead perhaps, further consider their values and give a well-thought out response to the PC question. This result suggests that the two-step telephone survey technique for eliciting PC responses may lead to a solution to the perplexing problem of ambivalence expressed by the respondents in Ready et al. (1995). This may allow those respondents who feel ambivalence to express it, without encouraging non-ambivalent respondents to shirk their duty to answer thoughtfully.

These results lead to opportunities for further research. First, more comparative studies should be conducted to determine the robustness of our results. Second, criterion validity experiments with real cash transactions could further improve

our understanding of the incentive compatibility of PC questions and the convergent validity of DC and PC valuation questions. Next, the additional information revealed by respondents with PC questions can be exploited to more efficiently estimate WTP with ordered regression techniques. Unlike the DC approach, however, where respondents only need to know whether their WTP is above or below a threshold, respondents to PC questions must consider an additional threshold and whether they are above or below it. Probabilistic PC questions may therefore elicit more information about the variance of individual WTP rather than information abut the intensity or certainty of preferences. One way to avoid the additional threshold is to use a continuous confidence scale after the DC response is obtained (Li and Mattsson 1995). Finally, comparisons between the discrete, probabilistic PC and continuous confidence scale approaches appears warranted.

Acknowledgements

This paper has benefitted from the helpful suggestions of Tim Haab, Rich O'Conor, Steve Swallow, Ken Wilson, participants in the East Carolina University Economics Seminar, participants in a session at the 1994 Southern Economics Association Meetings, and two anonymous journal referees. The data were collected with financial support from the East Carolina University Department of Economics and Survey Research Laboratory.

Notes

1. Johannesson et al. (1993) presented a PC question to respondents in a mail survey with five answers. After dropping the middle responses they group the data into yes and no categories and analyze the binary data and get theoretically valid results. No comparison between PC and DC data is made, however, since DC data was not collected. Li and Mattsson (1995) and Swallow et al. (1993) also collect PC data. After a DC question was asked, respondents were then asked how certain they were about their answer. Responses to the second question allows the econometric exploitation of the additional information about respondent preferences.
2. A copy of the survey questions is available upon request or can be obtained from URL: http://www1.ecu.edu/~ecwhiteh/data.htm.
3. Sensitive questions, such as household income, and difficult to answer questions, such as contingent valuations, tend to generate high item nonresponse rates which include protest and outlier responses. Lower education levels are related to item nonresponse on the CV question. Item nonresponse on the income question is correlated with certain demographic variables. Willingness to pay estimates should be expanded to the population with caution.
4. Substitute prices are also potentially important. However, with this data set all substitute prices attempted were either too collinear with the own-price or had no affect on the responses. For these reasons we dropped the substitute price from our analysis.
5. We do not consider the information treatments to be a construct validity test like those for own-price and income effects and WTP convergence. This investigation is more exploratory to see if information has the same effect in DC and sequential PC.
6. Poe et al. (1994) find that simulations approaches to constructing confidence intervals are too conservative. They acknowledge (footnote 1, p. 906) that the Cameron approach does not lead to this problem.

7. All models, WTP estimates, and standard errors for the Cameron WTP confidence intervals and t-tests are estimated with LIMDEP (Greene 1995). The data, LIMDEP program, and econometric output are available upon request or from URL: http://www1.ecu.edu/~ecwhiteh/data.htm.
8. The coefficient of variation is equal to the standard deviation divided by the mean WTP multiplied by 100.
9. The own-price variable is measured as each respondent's travel and time costs to the Pamlico Sound.
10. Other explanatory variables were included in exploratory regressions. Respondent concern and perceived effectiveness are both positively related to WTP while respondent knowledge has no effect on WTP. Since recreation participation is an endogenous variable, the exogenous own-price variable (which is negatively correlated with recreation participation) measures the effect of on-site resource use on WTP. None of this changes the results presented in this note, therefore we present the more parsimonious models.
11. The Cameron approach results in normally distributed estimates of mean WTP. Statistical comparisons of means can therefore be conducted using a straight forward t-test.

References

Arrow, Kenneth, Robert Solow, Edward Leamer, Paul Portney, Roy Radner, and Howard Schuman (1993), Report of the NOAA Panel on Contingent Valuation, *Federal Register* **58**, 4602–4614.

Bishop, George F. (1978), 'Experiments with the Middle Response Alternative in Survey Questions', *Public Opinion Quarterly* **51**, 220–232.

Blomquist, Glenn C. and John C. Whitehead (forthcoming), 'Resource Quality Information and Validity of Willingness to Pay in Contingent Valuation', *Resource and Energy Economics*.

Boyle, Kevin J. and Richard C. Bishop (1988), 'Welfare Measurements Using Contingent Valuation, A Comparison of Techniques', *American Journal of Agricultural Economics* **70**, 20–28.

Boyle, Kevin J., William H. Desvousges, F. Reed Johnson, Richard W. Dunford, and Sara Hudson (1994), 'An Investigation of Part-Whole Biases in Contingent Valuation Studies', *Journal of Environmental Economics and Management* **27**, 64–83.

Cameron, Trudy Ann (1988), 'A New Paradigm for Valuing Non-Market Goods Using Referendum Data: Maximum Likelihood Estimation by Censored Logistic Regression', *Journal of Environmental Economics and Management* **15**, 355–379.

Cameron, Trudy Ann (1991), 'Interval Estimates of Non-Market Resource Values from Referendum Contingent Valuation Surveys', *Land Economics* **67**, 413–421.

Cameron, Trudy Ann (1992), 'Combining Contingent Valuation and Travel Cost Data for the Valuation of Nonmarket Goods', *Land Economics* **68**, 302–317.

Carson, Richard T. and Robert Cameron Mitchell (1995), 'Sequencing and Nesting in Contingent Valuation Surveys', *Journal of Environmental Economics and Management* **28**, 155–173.

Cooper, Joseph C. (1994), 'A Comparison of Approaches to Calculating Confidence Intervals for Benefit Measures from Dichotomous Choice Contingent Valuation Surveys', *Land Economics* **70**, 111–122.

Freeman, A. Myrick, III (1993), *The Measurement of Environmental and Resource Values: Theory and Methods*, Resources for the Future, Washington, DC.

Greene, William H. (1995), *LIMDEP: Version 7.0 User's Manual*, Econometric Software, New York, NY.

Jakus, Paul M. (1995), 'Averting Behavior in the Presence of Public Spillovers: Household Control of Nuisance Pests', *Land Economics* **70**, 273–285.

Johannesson, Magnus, Per-Olov Johansson, Bengt Kriström, and Ulf-G. Gerdtham (1993), 'Willingness to Pay for Anti-Hypertensive Therapy – Further Results', *Journal of Health Economics* **12**, 95–108.

Kriström, Bengt (1993), 'Comparing Continuous and Discrete Contingent Valuation Questions', *Environmental and Resource Economics* **3**, 63–71.

Li, Chuan-Zhong and Leif Mattsson (1995), 'Discrete Choice under Preference Uncertainty: An Improved Structural Model for Contingent Valuation', *Journal of Environmental Economics and Management* **28**, 256–269.

Mitchell, Robert Cameron and Richard T. Carson (1989), *Using Surveys to Value Public Goods*, Resources for the Future, Washington, DC.

Poe, Gregory L., Eric K. Severance-Lossin and Michael P. Welsh (1994), 'Measuring the Difference of Simulated Distributions: A Convolutions Approach', *American Journal of Agricultural Economics* **76**, 904–915.

Ready, Richard C., John C. Whitehead, and Glenn C. Blomquist (1995), 'Contingent Valuation When Respondents are Ambivalent', *Journal of Environmental Economics and Management* **29**, 181–196.

Swallow, Stephen K., Thomas F. Weaver, and James J. Opaluch (1993), 'Ordered-Response Model for Ordinarily Dichotomous Referendum, Contingent Choice Data', Paper presented to the 1993 Annual Meeting of the American Economics Association, Orlando, Florida.

Whitehead, John C. (1995), 'Willingness to Pay for Quality Improvements: Comparative Statics and Theoretical Interpretations of Contingent Valuation Results', *Land Economics* **71**, 207–215.

Whitehead, John C. and Glenn C. Blomquist, Thomas J. Hoban, and William B. Clifford (1995), 'Assessing the Validity and Reliability of Contingent Values: A Comparison of On-Site Users, Off-Site Users, and Non-Users', *Journal of Environmental Economics and Management* **29**, 238–251.