View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

Buying Time: Real and Hypothetical Offers

V. Kerry Smith
Carol Mansfield

Discussion Paper 97-09

November 1996

Resources for the Future
1616 P Street, NW
Washington, DC 20036
Telephone 202-328-5000
Fax 202-939-3460

© 1996 Resources for the Future. All rights reserved.
No portion of this paper may be reproduced without permission of the authors.

Discussion papers are research materials circulated by their authors for purposes of information and discussion. They have not undergone formal peer review or the editorial treatment accorded RFF books and other publications.

Buying Time: Real and Hypothetical Offers

V. Kerry Smith and Carol Mansfield

Abstract

This paper provides the results of a field test of contingent valuation estimates within a willingness to accept framework. Using dichotomous choice questions in telephone-mail-telephone interviews, we compare responses to real and hypothetical offers to survey respondents for the opportunity to spend time in a second set of interviews on an undisclosed topic. Five hundred and forty people were randomly split between the real and hypothetical treatments. Our findings indicate no significant differences between people's choices with real and hypothetical offers. Choice models indicate the size of the offer and income were significant determinants of respondents' decisions, and these models were not significantly different between real and hypothetical offers.

Key Words: contingent value, real and hypothetical offer, willingness to accept, experiment

JEL Classification No(s).: C93, D12, Q2

Table of Contents

List of Figures and Tables

Buying Time: Real and Hypothetical Offers

V. Kerry Smith and Carol Mansfield*

I. INTRODUCTION

This paper reports the findings of a large scale field test of contingent valuation within a compensation (willingness to accept) setting.[1] Using a dichotomous choice framework, we offered the respondents to a telephone-mail-telephone (TMT) survey a payment to participate in a second TMT survey about an unspecified issue. Half of the sample was randomly assigned to a treatment offering real payments and the other half to one that described a plan where such payments "might" be offered. Our findings indicate that there are no significant differences between people's choices with real and hypothetical offers. Moreover, estimates of choice models based on these data indicate that income, the amount of the payment proposed or offered, each respondent's ability to answer the initial TMT survey, and the timing of those initial interviews (i.e., weekday versus weekend) were all significant determinants of respondents' decisions. Likelihood ratio tests suggest that none of these factors had a significantly different effect on the actual versus hypothetical intentions to participate in a subsequent TMT survey.

The motivation for this research stems from both the concerns raised with contingent valuation (Diamond and Hausman [1994]), as well as the diverse experimental evidence

---

* Arts and Sciences Professor of Environmental Economics, Duke University and Resources for the Future University Fellow; and Assistant Professor of Environmental Economics, Duke University respectively. Smith's research was partially supported by the UNC Sea Grant Program project number R/MRD-32. Thanks are due Richard Carson, Stanley Presser, and Pam Rathbun for comments or suggestions about related research and to Kris McGee for preparing several versions of this paper.

drawn from small, specialized samples involving purchases of private goods (Cummings et al. [1995]) and contributions for goods with public attributes (Duffield and Patterson [1992], Neill et al. [1994] and Brown et al. [1996]). The experiment described in this paper was designed to overcome some of the problems with past experiments comparing real and hypothetical responses to survey questions. We offered respondents a private good with which they had experience under circumstances that were less artificial than previous experiments.

The experimental design and framework used to analyze the data are described in the next section. Our results are summarized in Section III. In the fourth section, we discuss our findings and research design in more general terms in light of their implications for any attempt to design "controlled" experiments with "homegrown values."[2] The last section summarizes our conclusions.

II. EXPERIMENTAL DESIGN AND MODEL

A. Background

Past experimental evaluations of contingent valuation have been limited by several factors, perhaps most notably by their small samples and often specialized samples (e.g., undergraduates). Arguably experiments involving sales of pure private goods (e.g., strawberries, juicers, calculators, etc.) have realized the closest correspondence to the circumstances generally associated with actual choices. However, even in these cases the

---

[1] Bishop and Heberlein [1979] undertook the first simulated market experiment using a WTA perspective. Their sample involved 221 responses to the real component and 332 to the hypothetical component, making it about the same size as our sample. Subsequent experimental work has generally involved smaller samples.

context of the sale is not one respondents would typically encounter -- e.g., sales of strawberries door to door; solar calculators in a classroom; juicers in a church meeting hall; or sandwiches in a university food processing lab. Indeed, the process of recruiting of subjects alone creates an "experimental" setting to the proposed choices. Equally important, the distribution of prices must be truncated from above by the market price if the object offered is a well-known commodity. If the price is not truncated from above, then the experimental results can be questioned because the higher than "normal" level for the price itself may convey information to respondents about quality. This will be especially true if it falls outside what respondents perceive to be a "plausible price range" for this commodity.[3]

Experiments involving public or mixed public/private goods have been forced to use a donation as the payment vehicle and not a price. Donations do not convey the same financial consequences as prices. More specifically, the ordinary meaning of a donation implies "voluntary payment," while prices are not voluntary. Thus, none of the past experiments offer an ideal replication of an economic choice. As we describe below, our experiment is not free of these concerns either. Indeed, we will suggest this limitation is an inherent feature of any experiment involving homegrown values.

---

[2] Homegrown values refer to the value of a good to a subject that is independent of the value that an experimenter might "induce" for the objects of choice in the context of experimental economics. As Cummings et al. notes, "homegrown values are those that the subject brings to an experiment" (p. 260, note #2).

[3] This point relies on arguments for price as a signal of quality. See Akerlof [1970] and Wolinsky [1983] for early discussion. The need to design in a truncated range of price variation is due to Glenn Harrison in remarks at a conference, Camp Resources (an environmental economics research workshop sponsored by the Center for Environmental Resource Economics at Duke University), in Wilmington, N.C., 1995.

B. The Experiment

The subjects for the experiment were respondents who had just finished (at the time they were asked the question) a TMT survey about environmental issues in North Carolina. In the second telephone interview they were offered a specific dollar amount in either a real or hypothetical context to participate in a second TMT survey. Table 1 reproduces the text for the two formulations of our question. The timing and subject of the next survey were not identified. Each respondent was randomly assigned to one of the two treatments and offered one of five amounts (5, 10, 15, 25 or 50 dollars).[4]

Table 1: Text of Hypothetical and Real Offers*

Hypothetical - Plan for a Future Survey
• Researchers at Duke University **are considering** establishing a sample of households that would be contacted once a year to ask their opinions on programs like the ones I described. This is one of several plans. • If this approach were taken and if they could pay you [$ amount] for your continued involvement in two more interviews like the two we have done, would you participate? yes no
Real Offer for a Future Survey
• Researchers at Duke University **are establishing** a sample of households that would be contacted again to ask their opinions on programs like the ones I described. They **can pay** you [$ amount] for your **continued** involvement in two more interviews like the two we have done. They would send a check for [$ amount] to you when the next two interviews are finished. Will you participate in this new program? yes

[4] Due to survey budget constraints the probabilities of assigning respondents to each offer were not equal. The number of respondents assigned to each offer amount for the actual and hypothetical treatments are given as follows: (the percent of total sample for each treatment is in parentheses)

Offer	Actual	Hypothetical
5	59(23)	82(29)
10	76(29)	70(25)
15	72(28)	70(25)
25	27(10)	19(7)
50	27(10)	38(14)

There is no significant difference in the assignment patterns for the two cases. ($^2 = 6.68$ p-value = 0.15)

no (if yes) The researchers at Duke will send a confirmation letter about a month before the next set of interviews.

* The phrases in bold were displayed in bold to the telephone interviewers to indicate that they should be emphasized as part of the interview.

Our sample for the original TMT survey was derived from a random digit dialed (RDD) sample of households in North Carolina using Survey Sampling, Inc. for the initial set of telephone numbers. The initial RDD included 1,633 eligible telephone numbers. Of these, 405 refused to participate in the survey and 191 could not be contacted after 12 attempts, yielding a response rate of 62%.5 Eligible respondents were individuals 18 years or older who make decisions about the household budget. The survey was described as part of a research project for Duke University. If asked in the initial TMT survey, interviewers explained the purpose as "a study on what people think about issues that have been in the news recently."

---

5 Using the aggregate rates of agreement for the RDD sample assigned to each county along with Census information, it is possible to estimate a selection model. Because the data are aggregate we follow Papke and Wooldridge [1993], estimate a generalized linear model with a logistic link function and then apply Lee's [1983] proposed strategy for dealing with selection models based on non-normal distributions. The estimates are based on the McCullagh and Nelder [1989] GLIM framework with a binomial distribution with a logit link function using the aggregate acceptance rates by county for the eligible RDD numbers. The model is given as follows:

proportion agree to interview =	$-.045 \times 10^{-3}$ Median Household Income
	(-2.09)
	+.028 Proportion with College Degree
	(2.16)
	+.873 Number of People in Household
	(1.35)
	-.010 Proportion with High School Degree
	(-1.19)
	+.087 Intercept
	(1.78)
Number of Observations = 95	
Deviance =	134.33

This selection effect was not important to the choice models developed for our real/hypothetical comparison. It is not relevant to our case, both because of its insignificant effect, and more importantly, because our interest is in the population that completed both surveys. Nonetheless, it does offer for other TMT applications an alternative to the recent strategy suggested by Cameron et al. [1996] for mailed surveys.

The first interview collected attitude, knowledge, and demographic information from each respondent. It was conducted between October 26, 1995, and December 14, 1995. These initial interviews averaged approximately 14 minutes. While one contingent valuation question was asked of each respondent, the primary purpose of this first interview was to recruit respondents for the second interview about policies associated with two of three environmental issues assigned to each participant in the second interview.[6] 826 people agreed to complete the second interview and provided names and addresses. Of these 826 people, six had phones disconnected, and 540 of the remainder completed the second interview for a second stage response rate of 66%. These second interviews were conducted with the same individual who completed the first interview. Respondents were contacted for the second interview after they received a mailing describing two environmental problems and specific plans proposed to address them. These interviews were conducted between November 13, 1995, and January 23, 1996, and averaged about 12 minutes.

The sample for our experiment consisted of the 540 respondents who completed both parts of the TMT survey. The question given in Table 1 was the last one asked in the second interview. Table 2 summarizes a few demographic characteristics of the individuals randomly assigned to the real and hypothetical treatments. As expected, given the random assignment to the real and hypothetical treatments, there is no significant difference in the economic and demographic characteristics of respondents assigned to each group.

---

[6] The three issues were: controlling nutrient pollution in the rivers and estuaries of North Carolina by controlling the disposal of waste from large hog farms; beach re-nourishment projects along North Carolina's barrier islands where erosion has posed significant problems; and a lengthening of coastal outfalls for waste

water to enhance the ability of coastal communities to deal with deterioration of water quality in near-coastal areas.

Table 2: Characteristics of the Sample by Treatment[a]

Variable	Treatment	
	Real Offer	**Hypothetical Offer**
Household Income[b]	37,817 24633 [258] {29}	42,387 (29,938) [275] {28}
Age	47.3 (15.9) [260]	47.1 (14.9) [277]
Proportion female	.55 (.50) [261]	.54 (.50) [279]
Proportion white	.83 (.37) [261]	.85 (.35) [279]
Proportion retired	.21 (.41) [261]	.18 (.39) [279]
Proportion college graduates	.29 (.46) [261]	.36 (.48) [279]

a The table reports the arithmetic mean for each variable with the standard deviation in parenthesis below the mean. There are a few missing values for some variables. Sample sizes are reported in brackets below the standard deviation. Only in the case of household income are imputations used to replace the missing values. For income, the numbers below the sample size are the number of observations with missing income.

b Missing values for income are imputed based on the following equation:

household income = – 20571.54 + 298.63 Age
(–2.95) (3.86)
+ 17618.73 At Least College Graduate (=1)
(8.72)
+ 19737.89 Trade School Graduate (=1)
(2.26)
– 9476.41 Retired (=1)
(–2.71)
+ 17135.04 Married (=1)
(9.13)
–8791.68 Below High School Graduate (=1)
(–3.19)
+ 8803.03 Fully Employed (=1)
(3.80)
+ 6364.93 White (=1)
(2.85)
+ .89 Median Household Income Census for Respondent's County
(3.95)
– 6525.76 Unable to Answer Attitude and Preference Questions in First Interview.
(–2.71)
Number of Obs = 787
R^2 = .313.

The numbers in parentheses below the estimated coefficients are t-ratios for the null hypothesis of no association.

C. Modeling WTA Choices

All interviews require respondents to provide time, and presumably cognitive effort, to answer the questions posed of them. In the case of telephone interviews the "timing of that time", or the time of the week a respondent surveyed, is specific. While there is considerable information about the effects of monetary incentives on the response rates to mail surveys, much less is known about the incentive effects of monetary payments with TMT surveys.[7] In our model, each respondent's decision is assumed to result from evaluating a time allocation choice. Based on the wording of the offer, we assume respondents consider the financial payment offered, time required to complete both of their recent interviews (including the time they spent reading the mailed information), and any inconvenience caused by the specific timing of those interviews. While the questions in Table 1 provide no guarantees that the second TMT survey will be exactly the same as the surveys the respondents have just completed, this experience provides a reasonable basis for them to form a judgment about the proposed object of choice. Moreover, because the same source (Duke University) is being represented as involved with the new survey, this framing should reinforce these perceptions.

A simple economic model for the time allocation implied by agreeing to the interview parallels a labor supply or other time allocation decision (e.g., Deacon and Sonstelie [1985]). The respondent compares the offered price per unit for her time, p, with her reservation price

---

[7] See Gebler et al. [undated] for a meta analysis summarizing studies involving face-to-face, telephone and mixed model (i.e., face-to-face or telephone with mail-back returns) surveys. Incentives were found to increase response rates. They report that the studies in the literature for telephone surveys confirm that monetary incentives increase response rate and are statistically significant. Moreover, there is not a significant effect of prepayment versus promised payment across studies that compared them. The authors report that this may be due to variation across studies in the size of the monetary incentive. They describe one study they conducted holding the prepayment constant that found a higher response rate with prepayment.

for the anticipated time allocation, *r*. If $r > p$, she would refuse to participate, and if $r \leq p$, she would agree to participate.

To convert the offered payment in our survey to the unit price, *p*, we assume that the respondent uses the elapsed time of the two recent surveys as the basis for constructing *p* as a price per minute (e.g., offer relative to the total time for two interviews, ignoring time spent reading the mailed booklet). Because the amount of time spent responding to the survey may well be endogenous and determined by the saliency of the topic of the survey, we also consider empirical models with the choice based on the total offer alone.

These decisions also follow the conventional discrete choice framework used to analyze referendum style contingent valuation (CV) models where either a set of conditional indirect utility functions (Hanemann [1984]) or a variation function (Cameron [1988], McConnell [1990]) provides the behavioral model used to analyze respondents' stated choices to the CV questions. Both descriptions stem from a common constrained utility maximizing description of individual behavior and lead to an empirical model based on either the unit price or the total offer, *f*. The estimating equation treats the outcome of the decision process as a latent variable (either a utility difference in Hanemann's formulation or the difference between the reservation value and the offer in Cameron's model). The choice to participate or not is all that can be observed. Equation (1) outlines in general terms the framework based on the price per minute. A model based on the total offer would be the same as (1) with *p* replaced by *f*.

$$c = f(p, m, T, P, SE) \qquad (1)$$

where:

c = 1 if the respondent accepts the offer
0 otherwise

p = price per minute implied by the offer (f/t)
f = total dollar amount of the offer
t = respondent's perception of total time for the new TMT survey (assumed to equal respondent's actual elapsed time for the recently completed TMT survey)
m = respondent's household income
T = measure of convenience of the time (in our case whether interviewed on weekdays or weekends)
P = vector of other prices including respondent's wage rate
SE = vector of each respondent's socio-economic characteristics and attitudes

Differences in the assumed behavioral foundations for the choice will alter how equation (1), the choice model, is used to estimate "reservation prices" for TMT surveys. Moreover, following the literature analyzing choice data, we assume this reduced form model is not known by the analyst and includes an error so that the decision process leading to equation (1) is consistent with a probit (or logit) estimator.

III. RESULTS

Table 3 reports the contingency table describing the results of our real and hypothetical offers. The responses are pooled across the five different dollar offers. Our findings stand in sharp contrast with much of the earlier literature involving experimental comparisons of real and stated choices in that there is no significant difference between the choices made by respondents assigned to the two treatments.

One difficulty identified with some of this earlier literature arises from the absence of a significant price response for either the real or hypothetical offers (Smith [1994]). To investigate this possibility we estimated choice models based on these data in several different ways. In Table 4 we report six specifications for equation (1). An array of alternative models were considered. None changed the overall conclusions from those based on the choice functions reported in Table 4.

Table 3: Participation in TMT Survey With Real and Hypothetical Offers[a]

Variable Choice	Treatment	
	Real Offer	Hypothetical Offer
Refuse	44 (16.9)	51 (18.3)
Agree	217 (83.1)	228 (81.7)
Total	261 (100.0)	279 (100.0)
	^{2}(df=1) = .188, p-value = 0.665	
	Fisher's exact test p-value = 0.735	

[a] The numbers in the table report the count of each response for each type of offer. The numbers in parentheses are the percentage of the total responses for each type.

Income and the offer are statistically significant determinants of respondents' decisions. Their effects on respondents' choices generally agree with *a priori* expectations. For example, higher offers lead to greater prospects for participation. Income is also found to have a positive influence on participation. This finding may reflect the interest of higher income households in expressing preferences for publicly provided goods. A likelihood ratio test of the hypothesis that the real and hypothetical offers have the same choice functions cannot be rejected at any of the conventional significance levels. The p-values for the test for each alternative model specification are presented with the respective estimates for the model. Neither of these conclusions is altered by the offer price used -- whether per minute, *p*, or the total offer, *f*.

As noted in Table 2 (footnote b), we predicted income based on Census information and the demographic characteristics of those who did report household income. Results from the income imputation model are also reported in Table 2. Models (1) and (4) in Table 4 include the imputed value for income as well as a qualitative variable that identifies the

Table 4

Table 4 cont'd

observation as having a missing value. Models (2), (3), (5) and (6) include only those respondents who reported their incomes. The imputation process adds nearly fifty observations to our sample, but does not alter the qualitative conclusions with either the offer-per-minute or the total offer models (i.e., compare equations (1) and (2), and (4) and (5) in Table 4). While it is clear that replacing missing income values does affect the choice model, it does not alter the apparent correspondence between the models describing real and hypothetical offers.

The last set of results (models (3) and (6)) considers whether the selection effects associated with the completion of the second survey of the initial TMT influence our conclusion. Including an inverse Mills ratio, estimated based on the characteristics of the individuals who completed both the initial and follow-up surveys, does not alter our conclusions that real and hypothetical decisions can be described by the same choice model. The estimates for this model use the Huber [1967] covariance matrix to correct the probit estimates of the parameters' asymptotic standard errors for the heteroscedasticity induced by including an estimate of the inverse Mills ratio.[8]

The timing of the first set of interviews appears to have had a fairly complex effect on how people responded to the offers to participate in another survey (and this finding is the same for both the real and hypothetical groups). When both TMT interviews were conducted over weekends respondents had a clear aversion to participating in another TMT survey. Comparing the responses of those contacted twice on weekends with the rest of the sample, a

---

[8] White [1982] has demonstrated that this approach yields consistent estimates of the covariance matrix. This two step approach is a consistent, but not efficient method for dealing with selection effects with discrete choice models (see Zhang [1995]). Given the test results, it was not necessary to consider the use of bivariate probit for estimating the models

simple contingency table suggests that a significantly lower percentage (77.3% versus 83.8%) of these weekend people agree to participate.[9] This response was captured for our choice model in a variable, *weekday*, that equals one if at least one of the interviews was conducted on a weekday and zero otherwise. Thus, respondents who participated in at least one weekday interview, *ceteris paribus*, have a higher likelihood of agreeing to participate in a second TMT survey. This is reflected by the positive coefficient for the weekday effect variable.

However, an interaction term between the amount of the offer and the weekday variable suggests that these weekday respondents are less responsive to financial incentives than individuals who had both of their interviews take place on weekends. While this latter effect is not statistically significant in all models, the sign is consistent across all models and is consistently significant for models using the total offer (i.e. models (4) through (6)). Figure 1 illustrates how the composite of the two effects influences the index function for the probit model. Respondents interviewed on at least one weekday (WKD in Figure 1) have a higher intercept for the model used to predict the location parameter for the latent variable assumed to underlie choices in the probit framework, as indicated by the vertical difference, , between the lines WKD and WKND. However, the slope of WKND is steeper than WKD, implying that financial incentives have a larger effect on the likelihood of participating within the group of respondents also answered both surveys on weekends.

---

[9] Rejection of the null hypothesis was at a p-value of .067 with a Fisher exact test.

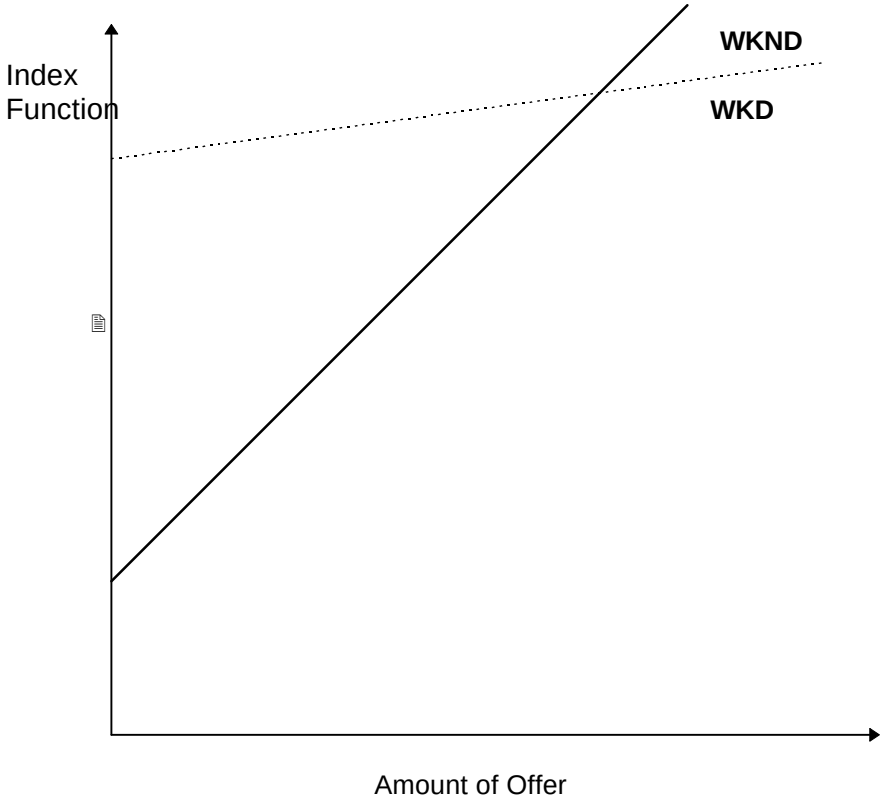

Figure 1: Illustration of the Weekday versus Weekend Effect

Because this issue has implications for indirect approaches to non-market valuation requiring measures of the opportunity cost of time (e.g. the travel cost recreation demand model), we attempted to estimate separate effects for the timing of the first interview (weekday or weekend), the number of days between interviews, as well as a variety of other alternative specifications for the days of the week associated with the interviews. The timing of the initial contact, weekday or weekend, does appear to reduce the prospects for agreeing to another interview, but this effect was not statistically significant.

To our knowledge the literature on monetary incentives in surveys has not considered TMT interviews. In another context, Deacon and Sonstelie [1985] found a lower opportunity cost of time on weekends using estimates from a natural experiment involving time spent waiting in line for low-priced gasoline (versus purchasing at a higher price without the wait) to compute the opportunity costs of time. Their findings imply estimates for the value of time that are 8.9 percent lower for weekend in comparison to weekday waiting time.

There is also no specific information on weekday versus weekend effects in past surveys. Some analysts have suggested that the "folklore" of political polling using telephone surveys warns against weekend interviews.[10]

Thus, at a general level, our models for respondents' decisions about an additional TMT survey do support the conclusion that an economic tradeoff appears to have influenced in both the real and hypothetical choices. Nonetheless it is important to also acknowledge that the financial consequences may not be as apparent as they are in a willingness-to-pay setting. Consequently, in computing the implied opportunity costs of time we considered

[10] Richard Carson identified this guidance as part of the practice of political polling in the 1980s.

only the estimated reservation prices for individuals where financial incentives seemed to offer the prospect for changing decisions (i.e., those who had both interviews on weekends). The results for those respondents with at least one weekday interview are largely unresponsive to financial incentives and therefore imply large reservation prices for time required for another TMT survey. We conclude that they are unlikely to be informative about these respondents' opportunity costs of time, generally.

Models based on the offer per minute and the total offer (with and without income imputation) were used for those interviewed on weekends. While these estimates are random variables, simple comparisons with other measures of the opportunity cost of time suggest considerable agreement with other approaches for estimating the value of time and with what has been found in previous studies in the literature. For example, a common practice in the literature is to use income per hour as a proxy for the hourly wage in estimating the opportunity cost of time.[11] Using household income and 2000 hours (as the work time) for the respondents to our offers, we have estimates of $20.66 per hour for the mean ($19.28 as the median) "hourly wage" in 1995 dollars. Based on the choice equation with the imputed income measure for households who did not report their income, the average hourly reservation value of time was estimated at $21.18 (for the model based on offer per minute) and $19.65 using the total offer. Models estimated from samples dropping the respondents with missing income implied higher reservation values (averages of $32.16 and $32.34 respectively).

---

[11] See Bockstael, Strand and Hanemann [1987] for discussion of alternative approaches for dealing with the opportunity cost of time in travel cost recreation demand models.

Table 5 compares our estimates with those from Deacon and Sonstelie by income class as well as the "income per hour" based on our respondents' reported household income for 1994 (all reported in 1995 dollars). Our estimates of the mean reservation value of time generally exceed their upper bound for comparable income classes. If we consider the general range of estimates between the two studies there is some overlap in the implied valuations, with our findings implying greater values for time.

Table 5: Hourly Reservation Prices by Household Income: A Comparison[a]

Income Class	TMT Choices		Household Income per Hour	Deacon - Sonstelie	
	Offer per Minute	Total Offer		Lower Bound	Upper Bound
30,000 - 45,000	26.22	24.30	19.28[b]		
45,000 - 60,000	35.64	33.51	26.99	16.52	25.34[c]
60,000 - 75,000	45.06	42.64	34.71		
75,000 - 90,000	54.48	51.77	42.42	20.82	31.92
90,000 - 100,000	62.28	59.38	48.85		
Over 100,000	92.04	88.26	73.24		

[a] The estimates for the TMT choices are based on the estimates from the offer per minute and total offer models replacing missing income values with imputed income. They consider respondents interviewed on weekends and evaluate each model's implied reservation price for each sub-sample.

[b] These estimates are converted to 1995 using the CPI for 1994 and 1995, the year which the household income was requested, implying about 28% increase in the general price level. (e.g., 148.2 to 152.4).

[c] These estimates are taken from Deacon and Sonstelie [1985] Table 4. They relate to values for fully employed respondents. The income classes from Deacon and Sonstelie were defined for individuals while ours relate to households. Their classes were 30,000 to 40,000 and over 40,000. We adjusted these using the CPI for 1980 in a 1982 - 84 base (82.4) and 1995 in the same base (152.4) to convert the income range to approximately match ours. However, because our measure is for the household their categories remain mismatched (and likely relate to higher income ranges at the household level). Their estimates relate to weekday tradeoffs. We did not apply the 8.9% reduction they found because it would affect the lower and upper differently than a simple scaling. The income range was adjusted to 1994 (the year requested for income) and the reservation price to 1995 by the relevant CPI.

These results stand in sharp contrast with recent estimates reported by Hausman et al. [1995] in the context of a model choice decision for recreation fishing trips to Alaska. Adjusting for the timing of these choices in relation to our sample, their results would imply

an opportunity cost (or reservation value) less than one half our lowest estimate. While this seems quite disparate, it is not necessarily inconsistent with our higher values for time. This conclusion follows because their choice model reflects a different decision margin. Recreationists in their survey have already decided to engage in recreation and the decision is among travel modes to reach the recreation site, considering both time and cost. Thus, the opportunity cost is the foregone time that could be spent in the activity as opposed to traveling to it. Furthermore, the travel itself may well be an important positive component of the overall experience in the context of fishing in Alaska. The distinction between undertaking a trip and incurring all of its costs, as opposed to re-allocating expenditures to be made for a trip among alternative uses is analogous to the difference in expenditure models used to define compensating variation and a conditional compensating variation (see Hanemann and Morey [1992]). Our experiment involves a decision of whether to engage in an activity (i.e., the second TMT survey) at all. By contrast their analysis is more likely to involve a re-allocation of the total recreation time between travel and on-site activities. It is therefore a different choice margin. Moreover with a second TMT survey, respondents who agreed to participate in the activity recognize the time required displaces leisure time. It is not re-allocation of the total time used for the recreation experience.[12] Thus, we conclude that for respondents on weekends, the decisions about real and about hypothetical offers imply plausible reservation values for time.

---

[12] There is another aspect that distinguishes participation in a survey from a decision about mode of travel in a recreation trip. The time allocation choices for a survey can serve a signaling role. Some time ago Spence [1973] offered a very interesting discussion of how time can serve a different signaling role than money. One direct implication of his arguments is that if our respondents expect the survey will deal with environmental issues (because of the content of the initial TMT surveys) and they wish to signal greater concern, they may well agree at a higher rate, implying our model would understate their "true" opportunity costs of time.

IV. DISCUSSION

The challenge of simulating "real" economic choices in an experimental setting as part of any comparison of actual and hypothetical decisions is more formidable than the literature to date has fully acknowledged. For experiments involving private goods, the process of recruiting subjects and specifying the prices and rules for exchange, along with other specifics of the experiment should be expected to influence the choices that respondents make.[13] Unfortunately, we do not yet know and may never know whether the experimental conditions influence "real" and "hypothetical" choices in an identical manner -- and the problems are even more difficult when the experiment involves a public good. This complicates our ability to compare results across experiments and ultimately to determine which set of experimental results is "right."

Our results offer a more encouraging picture of the ability of CV surveys to produce reliable data than some other experimental tests, at least in the context of a private good. Why do our results differ from these past experiments? In our experiment, the "good" was a private good with which the respondents had experience having just completed a survey. Furthermore, the circumstances of the choice were much less artificial than previous experiments, especially those conducted in laboratories using undergraduate subjects. Marketing firms routinely solicit subjects through telephone interviews. Moreover, telephone surveys are not exotic commodities. In fact, most people are aware that public opinion polling takes place all of the

---

[13] For example in the Shogren et al. [1994] study of food contamination risks, respondents were required to consume the product before leaving the location of the experiment. Other comparisons of real and hypothetical responses sold a framed water color of a southwest rural scene by a Navajo artist, a framed print of a 16th century map of the world (see Neill et al. [1994]), and calculators in meeting rooms for undergraduate classes.

time, even if chances are low that a particular individual has actually participated in an opinion poll.

The Cummings et al. [1995] experiments assure respondents to the hypothetical survey that it is not an "opportunity to buy" the product. Rather it is an interest (presumably on the part of the researcher conducting the experiment) in respondents' answers to a question -- "Would you be willing to pay $X for the juicer?" To an economist, this is the same as a purchase question. For a respondent it may not be. By contrast, our approach uses a proposed or potential program and asks "if" it were available and "if" the Duke researchers "could pay you $X would you participate?" Framing CV questions using the context of a plan to accomplish a stated objective is somewhat more consistent with the NOAA Panel's (Arrow et al. [1993]) recommendations. A similar framework was used by Carson et al. [1996] in surveys that have found support for willingness to pay (WTP) estimates derived from contingent valuation. Nonetheless, there is no theoretical evidence that assures this approach to framing the CV question is superior.

While the object of choice and circumstances of choice in our experiment may seem less contrived than typical experiments, the survey is not free from difficulties. Respondents to our questions could have interpreted the questions in one of three ways: (1) respondents in the real and hypothetical treatments understood their respective questions, (2) both sets of respondents thought the questions were real, or (3) both sets of respondents thought the questions were hypothetical.[14]

Electric juicers were sold at church (without claiming to represent a charity) (Cummings et al. [1995]), and strawberries door-to-door (Dickie et al. [1987]).

[14] It may seem that we could have developed more information to discriminate among these explanations by asking more questions of our respondents. Because this was the last question of the second interview further

Given the close correspondence between the responses of the two groups, the three alternative reactions to our questions have different implications for the interpretation of our results. If both groups of respondents understood their questions, we have demonstrated that at least in this circumstance, the individuals responding to the hypothetical question knew the question was hypothetical but gave the same response that they would have given to a real question.

An alternative explanation of our results holds that either both groups considered the questions to be real or both groups considered the questions to be hypothetical. Either of these interpretations would also explain the similarity of the responses between the two groups. Which of these alternatives is most likely? Consider the circumstances: all the respondents have just finished a TMT survey in which the two telephone calls were separated by an average of 26 days. Our analysis indicates that both price and income, in addition to other variables, are significant, and the responses imply plausible estimates for the value of a respondent's time. From this evidence, we consider it much more likely that if respondents misinterpreted the questions, all the respondents interpreted the two questions as real, as opposed to hypothetical.

If all the respondents believed that the questions were real, this supports the recommendations of the NOAA panel. Despite qualifications, people do not notice the conditionality of the hypothetical questions and answer as they would to a real choice,

---

inquiries about what respondents thought when answering these questions were not included. Even if they were, there is substantial evidence in the literature that further questions asking respondents to explain why they answered a question in a particular way provide limited information about their actual motives. Based on the Krosnick and Fabrigar's [1992] review, respondents seem unlikely to offer clear-cut answers to these questions.

because they think it is a real choice. This interpretation would suggest that our framing of the CV question is effective, if unethical.

Past evidence with hypothetical WTA based CV surveys indicate a clear record of difficulties in using this type of question format. The WTA questions often generate implausibly large monetary values, as well as relatively little sensitivity of stated choices to the payment offered or to household income. They are often insensitive to attitudes that one would hypothesize should be related to the decisions involved. Indeed, this record is cited in the NOAA Panel's recommendations against a WTA format. Of course, we should also recognize what we observed earlier. In this survey the offer to purchase time is likely to be familiar to these respondents. Most WTA studies have used objects of choice where either the offer was viewed illegitimate or acceptance of the offer could be regarded as inappropriate by a reasonable fraction of a representative sample (e.g., willingness to accept health risks or compensation for the destruction of specific wilderness areas). Even Bishop and Heberlein's WTA dichotomous choice experiments for well recognized private commodities (hunting permits) displayed fairly large numerical differences between the reservation values implied by the responses to the hypothetical and simulated offers. For the goose permits it was $101 versus $63 (hypothetical /real) and for deer permits $420 versus $153.

In contrast, as we noted earlier, our estimates do display significant effects for the offer amount and household income. For the subset of our sample more willing to consider financial incentives, the models' estimates for the value of their time display reasonably close correspondence to updated estimates from literature based on actual choices. Finally, the parameter estimates from the models for hypothetical and real offers were not significantly

different, and the offer interaction coefficient (i.e., a qualitative variable for real versus hypothetical offers and our measures of the dollar amount of the offer) was not significantly different from zero. Thus, we think it is unlikely that our results are due to respondents considering both question framings as hypothetical.

On a separate issue, one might argue that respondents perceive the financial consequences of a WTA choice differently than a WTP choice: WTA does not function as an economic tradeoff because the individual does not make a payment. "Money in hand" (i.e., income) is not allocated in response to the choice, as it would be with experiments using a WTP perspective. As a consequence, the real/hypothetical distinction within a WTA survey is not informative about CV's performance in other contexts.

This reasoning would imply that the amount of the offers and a respondent's household income should not affect choices, and the results clearly contradict this hypothesis. In fact, there are economic tradeoffs associated with payment offers. The respondent who refuses to participate in a second survey in the real offer treatment foregoes money that could be used to purchase something else of value. Since our statistical models suggest that both the level of an individual's household income and the size of the offer are important, we conclude a genuine tradeoff is present.

Because the payment is not immediate in our case, it might be argued that this separation between choice and compensation provides another potential reason why respondents might incompletely perceive the financial consequences of the offer. Here again, we would argue the answer is ambiguous for several reasons. Respondents have experienced the reinforcement of two interviews separated by an average of about 26 days. As a result,

they are aware of time lapses and yet fulfillment of stated intentions to interview.[15] This separation is not unique to our experiment. As Faith and Tollison [1981] have suggested, *ex post* payment is a rational response in designing institutions (or rules) to control significant transaction costs where there are large interpersonal differences in the information available to buyers and sellers. For example, people are recruited to participate in marketing experiments and focus groups with the promise of payment after participation. The situation posed here is comparable. Thus, this line of reasoning cannot explain between the real and hypothetical responses observed in this experiment.

V. CONCLUSION

Overall we believe our findings offer quite a different perspective on the ability of contingent valuation to mimic real choices with private goods. The pool of subjects was familiar with the object of choice, and the terms of payment were consistent with real world contracts for these types of services. People encounter opportunities to participate in mall intercept surveys and focus groups with delayed payment. They are offered chances to collect coupons, proof of purchase symbols, and other materials for monetary rebates that often require 6 to 8 weeks of waiting time. In each case a choice is made and effort is expended before payment is made.

Both the real and the hypothetical responses in our experiment were sensitive to income and to the offer amount. In addition, the estimates implied for the values of time for those respondents displaying responsiveness to financial incentives fall within the same

---

[15] Funding for a follow-up project was eliminated in the cutbacks imposed by NOAA in the 1995-96 budget recession. Efforts are underway to fund recontacting these households with a follow-up TMT survey. Nonetheless, the interviews and payment are not completed at this time.

general range as other studies' measures of the value of time. The choices offered to over 500 respondents in our experiment are different than what has been offered in past experiments but not obviously biased to create consistency between real and hypothetical offers. Indeed, they conform to the NOAA Panel's recommendation to use some type of plan to describe the object of choice in CV surveys.

Nonetheless, having stated a rather strong conclusion it would not be prudent for us to ignore an important limitation in our experiment. We cannot determine the "realness" to our sample respondents of either our actual offer to participate in a second TMT survey and be paid or the degree of "hypotheticality" of our proposed plan for such an offer. However, this limitation is not unique to our experiment. It is inherent in all experiments with homegrown preferences. Unlike conventional experiments, where the analyst can induce preferences with a monetary reward system and limit the extent of respondent and context effects, efforts to evaluate "real" choices must confront the difficulties in "simulating" any actual choice. Economists simply do not know how the circumstances of each choice influence the reliability of information from comparisons between simulated and stated choices. This will remain a limitation in using experiments as a basis for evaluating the correspondence between actual and stated choices. We believe these concerns are comparable to those raised some time ago with respect to comparisons between travel cost (or hedonic models) estimates and CV results. The primary difference between the two types of comparisons (i.e. "indirect methods versus CV" and "simulated market versus CV") arises in the error propagating model assumed to be relevant for each type of comparison. With this generic qualification to all experimental comparisons, we conclude that there do not appear to be reasons specific to our

application that would make it more likely to display consistency between real and hypothetical choices. Our empirical findings across three different types of comparisons -- contingency tables, choice models, and imputed values of time in relation to the literature -- provide clear evidence of consistency between actual and stated choices.

REFERENCES

Akerlof, G. A. 1970. "The Market for Lemons: Qualitative Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, vol. 84 (December), pp. 488-500.

Arrow, K., R. Solow, P. R. Portney, E. E. Leamer, R. Radner, and H. Schuman. 1993. "Report of the NOAA Panel on Contingent Valuation," *Federal Register*, vol. 58, no. 10, January 15, pp. 4601-4614.

Bishop, Richard C., and Thomas A. Heberlein. 1979. "Measuring Values of Extramarket Goods: Are indirect Measures Biased?" *American Journal of Agricultural Economics*, vol. 61 (December), pp. 926-930.

Bishop, Richard C., and Thomas A. Heberlein. 1990. "The Contingent Valuation Method" in Rebecca L. Johnson and Gary V. Johnson, eds., *Economic Valuation of Natural Resources: Issues, Theory and Application* (Boulder, Colo., Westview Press), pp. 81-104.

Bockstael, Nancy E., Ivar E. Strand, and W. Michael Hanemann. 1987. "Time and The Recreational Demand Model," *American Journal of Agricultural Economics*, vol. 69 (May), pp. 293-302.

Brown, T. C., P. A. Champ, R. C. Bishop, and D. W. McCollum. 1996. "Response Formats and Public Good Donations," *Land Economics*, vol. 72 (May), pp. 152-166.

Cameron, Trudy Ann. 1988. "A New Paradigm for Valuing Non-Market Goods Using Referendum Data," *Journal of Environmental Economics and Management*, vol. 15 (Sept.), pp. 355-379.

Cameron, Trudy Ann, W. Douglass Shaw, Shannon R. Rogland, Sally Keefe, and John M. Callaway. 1996. "A Method for Nonresponse Correction the Analysis of Mail Survey Data." Paper presented at W133 Meetings, Jekyl Island, Georgia.

Carson, R. T., R. C. Mitchell, W. M. Hanemann, R. J. Kopp, J. A. Krosnick, S. Presser, P. A. Ruud, and V. K. Smith, with M. Conaway and K. Martin. 1996. "Was The NOAA Panel

Correct About Contingent Valuation?" Quality of the Environment Division Discussion Paper No. 96-20, Resources for the Future, Washington, D.C. (May).

Cummings, R. G., G. W. Harrison, and E. E. Rutström. 1995. "Homegrown Values and Hypothetical Surveys: Is the Dichotomous Choice Approach Incentive Compatible?" *American Economic Review*, vol. 80 (March), pp. 260-266.

Deacon, Robert T., and Jon Sonstelie. 1985. "Rationing By Waiting and the Value of Time: Results from a Natural Experiment," *Journal of Political Economy*, vol. 93 (December), pp. 627-647.

Diamond, Peter A., and Jerry A. Hausman. 1994. "Contingent Valuation: Is Some Number Better Than No Number," *Journal of Economic Perspectives*, vol. 8 (Fall), pp. 45-64.

Dickie, Mark, A. Fisher, and S. Gerking. 1987. "Market Transactions and Hypothetical Demand Data: A Comparative Study," *Journal of the American Statistical Association*, vol. 82 (March), pp. 69-75.

Duffield, John W., and David A. Patterson. 1992. "Field Testing Existent Values: An Instream Flow Trust Fund for Mountain Rivers." Paper presented to W-133 Meetings, 14th Interim Report edited by R. B. Rettig, Department of Agricultural and Resource Economics, Oregon State University.

Faith, Roger L., and Robert D. Tollison. 1981. "Contractual Exchange and the Timing of Payment," *Journal of Economic Behavior and Organization*, vol. 1 (Dec.), pp. 325-342.

Gebler, Nancy, Eleanor Singer, Trivellore Raghunathan, and Kate McGonagle. Undated. "The Effects of Incentives for Response Rates in Face-to-Face, Telephone and Mixed Mode Surveys." Survey Methodology Program, Survey Research Center, University of Michigan.

Hanemann, W. Michael. 1984. "Welfare Evaluations in Contingent Valuation Experiments With Discrete Responses," *American Journal of Agriculture Economics*, vol. 66 (August), pp. 332-341.

Hanemann, W. Michael. 1994. "Valuing the Environment through Contingent Valuation," *Journal of Economic Perspectives*, vol. 8 (Fall), pp. 19-44.

Hanemann, W. Michael, and Edward R. Morey. 1992. "Separability, Partial Demand Systems, and Consumer's Surplus Measures," *Journal of Environmental Economics and Management*, vol. 22 (May), pp. 241-258.

Hausman, Jerry A., Gregory K. Leonard, and Daniel McFadden. 1995. "A Utility-Consistent, Combined Discrete Choice and Count Data Model: Assessing Recreational Use Losses Due to Natural Resource Damage," *Journal of Public Economics*, vol. 56 (January), pp. 1-30.

Heckman, James. 1979. "Sample Bias as a Specification Error," *Econometrica*, vol. 47 (January), pp. 153-162.

Huber, P. J. 1967. "The Behavior of Maximum Likelihood Estimates Under Non-Standard Conditions," in *Proceedings of the Fifth Symposium On Mathematical Statistics and Probability* (Berkeley, Calif., University of California Press).

Krosnick, Jon A., and Leandre R. Fabrigar. 1992. "'Asking Why?': The Validity of Introspection." Unpublished paper, Department of Psychology, Ohio State University, March.

Lee, Lung-Fei. 1983. "Generalized Econometric Models With Selectivity," *Econometrica*, vol. 51 (March), pp. 507-512.

McConnell, K. E. 1990. "Models for Referendum Data: The Structure of Discrete Choice Models for Contingent Valuation," *Journal of Environmental Economics and Management*, (January), pp. 19-34.

McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*, second edition (London, England, Chapman and Hall).

McFadden, Daniel. 1974. "The Measurement of Urban Travel Demand," *Journal of Public Economics*, vol. 3 (November), pp. 303-328.

Neill, H. R., R. G. Cummings, P. T. Ganderton, G. W. Harrison, and T. McGucken. 1994. "Hypothetical Surveys and Real Economic Commitments," *Land Economics*, vol. 70 (May), pp. 145-154.

Papke, Leslie E., and Jeffrey M. Woolridge. 1983. "Econometric Methods for Fractional Response Variables With An Application to 401(k) Plan Participation Rates." NBER Technical Paper #147 (November).

Shogren, Jason F., S. Y. Shin, Dermot J. Hayes, and J. B. Kliebenstein. 1994. "Resolving Differences in Willingness to Pay and Willingness to Accept," *American Economic Review*, vol. 85 (March), pp. 260-266.

Smith, V. Kerry. 1994. "Lightening Rods, Dart Boards, and Contingent Valuation," *Natural Resources Journal*, vol. 34, no. 1, pp. 121-152.

Spence, A. Michael. 1973. "Time and Communication In Economic and Social Interaction," *The Quarterly Journal of Economics*, vol. 87, pp. 651-667.

White, Halbert. 1982. "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, vol. 50 (January), pp. 1-25.

Wolinksy, Asher. 1983. "Prices as Signals of Product Quality," *Review of Economic Studies*, vol. L (October), pp. 647-658.

Zhang, Xiaolong. 1995. "Integrating Resource Types, Payment Methods, and Access Conditions to Model Use and Nonuse Values: Valuing Marine Debris Control," unpublished Ph.D. thesis, North Carolina State University, Raleigh, N.C.