An Experimental Investigation of Explanations for Inconsistencies in Responses to Second Offers in Double Referenda

Anthony C. Burton †
Economics and Operational Research Division, Department of Health
Quarry House, Quarry Hill, Leeds, UK, LS2 7UE

Katherine S. Carson†
Department of Economics and Geography, United States Air Force Academy
HQ USAFA/DFEG, 2354 Fairchild Drive, Ste. 6K110, USAF Academy, CO 80840-6299

Susan M. Chilton†
University of Newcastle upon Tyne Business School-Economic
Newcastle upon Tyne, UK, NE1 7RU

W. George Hutchinson†
Department of Agricultural and Food Economics, Queen's University
Belfast, Northern Ireland, UK, BT9 5PX

Presented at the Annual Conference of the Royal Economic Society, Swansea, April 2004

Suggested Running Head: RESPONSES TO SECOND OFFERS IN DOUBLE REFERENDA

Please direct all correspondence regarding this manuscript to:
Susan M. Chilton
University of Newcastle upon Tyne Business School-Economics
Ridley Building
Queen Victoria Road
Newcastle upon Tyne
NE1 7RU
PH: +44 (0)191 2226833
FAX: +44 (0)191 2225780
Email: s.m.chilton@ncl.ac.uk

*This research is funded by the National Science Foundation (Division of Social and Economic Sciences, Decision, Risk, and Management Sciences Program, #SES-9986407). All opinions herein are solely those of the authors and do not reflect the policy either of the National Science Foundation or of the authors' respective institutions. Tony Burton wishes to thank the Northern and Yorkshire Regional Health Authority and Brian Ferguson for allowing him the time to conduct this research. Any remaining errors or omissions remain the responsibility of the authors.
bb† The convention of alphabetical authorship is adopted. No seniority implied.

ABSTRACT

This paper demonstrates the potential for induced preference experiments to test previously unverified explanations of observed behavior in contingent valuation surveys. The NOAA Panel on Contingent Valuation called for experimental evidence on potential biases in the double referendum format. We test Carson, Groves, and Machina's [6] simple cost uncertainty and weighted averaging explanations of inconsistent responses to follow-up offers in such double referenda against a baseline of certainty and truthful preference revelation. The results find evidence to support the weighted average hypothesis. Results regarding the cost uncertainty hypothesis are more ambiguous and merit further investigation.

KEY WORDS:
Experimental Economics
Contingent Valuation
Double Referenda
Induced Preferences

AN EXPERIMENTAL INVESTIGATION OF EXPLANATIONS FOR INCONSISTENCIES IN RESPONSES TO SECOND OFFERS IN DOUBLE REFERENDA

1. INTRODUCTION

In 1993, the NOAA Panel on Contingent Valuation recommended the use of the hypothetical referendum question format in surveys designed to estimate losses associated with natural resource damage. A hypothetical referendum question asks, for example, "If your cost for program A were $D per year, would you vote for program A?" The panel recommended the hypothetical referendum format, because,

> ...as far as strategic reasons go, a respondent who would not be willing to pay D dollars has no reason to answer "Yes," and a respondent who would be willing to pay D dollars has no reason to answer "No." ([3], p. 4606)

Practitioners of contingent valuation often employ double referendum questions (frequently referred to as double bounded dichotomous choice formats) to generate estimates for willingness to pay. This form of referendum requires the respondent to answer two 'YES-NO' questions of the type: "Would you pay $D for A?" The first question poses a value of $D previously selected from a range of values. The second question poses a value taken from a sub-range (either higher or lower than $D) that depends on the respondent's answer to the first question. Arguments favoring the use of the double referendum format in surveys center on its simplicity and the substantial gains in statistical efficiency compared to asking only one question [10].

An important stylized fact regarding double referendum surveys is that the mean or median willingness to pay estimated using responses to the first valuation question often differs from that estimated using responses to both the first and second questions. Furthermore, the error terms of

the two estimated valuations are not perfectly correlated. The number of negative responses to the second question is often higher than would be expected from the distribution of values based on responses to the first question alone [10, 12]. A higher than expected frequency of Yes-Yes and No-No responses has also been observed [5]. A number of econometric methods to address this issue have been proposed (e.g. [2, 5]). Several theoretical explanations are advanced for respondents' inconsistent responses to the first and second referendum questions in DeShazo [9] and Carson, Groves, and Machina [6], hereafter CGM.

Part of the Blue Ribbon Panel's recommendations about contingent valuation includes an admonishment about the use of survey formats that deviate from the simple single referendum format. In its *Guidelines for Value Elicitation Surveys,* the panel states, "If a double-bounded dichotomous choice or some other question form is used in order to obtain more information per respondent, experiments should be developed to investigate biases that may be introduced " ([3], p. 4612). Given the inconsistencies in WTP estimates based on the first and second responses described above and the concerns of the NOAA Panel regarding double-bounded dichotomous choice surveys, it is interesting that there has been no experimental research on the topic to date. The research to date has focused on single referenda [4, 7, 13]. While the results of these studies diverge somewhat, it seems timely to begin a parallel research agenda centered on double referenda. Indeed, it is likely that results from such studies, depending on their design, can provide further evidence on responses to single referendum questions as well.

In this paper, we report the results of a set of experiments designed to test two of the four theoretical explanations proposed by CGM to explain the stylized facts about double referendum surveys. We term these two theoretical explanations the Weighted Average hypothesis and the Uncertain Cost hypothesis. The purpose of these experiments is to determine if subject behavior

is consistent with the theoretical predictions when the incentives described in the theories are induced, consistent with the use of experiments as a means of theory falsification (Davis and Holt [8], p. 32). Our purpose is not to replicate a double bounded dichotomous choice field survey in the laboratory. Instead, using a single referendum question we replicate and study the effect of the types of uncertainty that CGM suggest occur when subjects respond to the second round questions in a double referendum. This allows us to isolate and test whether CGM's theoretical explanations, when induced in a laboratory setting, will produce subject responses which are consistent with their predictions and are capable of explaining observed behavior in the second round of a double referendum.

Our results provide support for the CGM Weighted Average hypothesis. However, the results do not provide clear evidence regarding the Uncertain Cost hypothesis. While our findings do provide some insight into possible explanations of subject behavior when responding to double referendum surveys, we do not claim to provide definitive evidence regarding which type of response is actually driving behavior observed in the field. This is perhaps an aim for future experiments. Thus, the paper should be seen in the wider context of demonstrating how induced preference experiments can be used to aid understanding of the CV process. Such experiments could be used to test new theoretical explanations of subject behavior and/or to aid in the development of new empirical methods to obtain improved willingness to pay estimates from field survey data.

The remainder of the paper is organized as follows. Section 2 contains a description of the theories posed by CGM to explain observed responses to double bounded dichotomous choice surveys. The experimental design is described in Section 3. We report the results in Section 4 and present some concluding thoughts in Section 5.

2. THEORETICAL BACKGROUND

CGM hypothesize about possible subject behaviors that may explain the stylized facts associated with double-bounded dichotomous choice contingent valuation questions. It is important to note that when willingness to pay distributions are estimated, it is assumed that subjects are not responding strategically to the valuation questions. In addition, it is assumed that subjects accept the costs presented to them as certain. Therefore, the estimated distributions of willingness to pay are generated under the assumptions of certainty and truthful preference revelation. A violation of either of these assumptions could generate the empirical irregularities observed in contingent valuation data. All of the hypotheses in CGM address the issue of subject uncertainty about cost. This uncertainty stems from the fact that subjects are asked two willingness to pay questions during the course of the survey. The nature of subject uncertainty determines the optimal response. CGM hypothesize that there are four possible effects of the second valuation question, which we will term the Uncertain Cost, Weighted Average, Bargaining, and Quantity/Quality Shift hypotheses. These four hypotheses are described in turn below.

In the Uncertain Cost hypothesis, subjects respond to the uncertainty created by the second valuation question by treating the second question independently from the first, but treating the new cost as an uncertain value with a mean equal to the second stated cost amount. Thus, respondents' risk preferences play a role in determining their answers to the second question. CGM make an additional assumption that subjects are risk averse. As a result, the distribution of stated willingness to pay will be skewed to the left as more subjects respond "No" to the second question than if the cost were known with certainty.

In the Weighted Average hypothesis, subjects respond to the second cost amount by assuming that the true cost is 'somewhere in the middle' of the two costs suggested by the survey administrator. As a result, subjects answer the follow up question as if the cost were a weighted average of the first and second costs, where the weights sum to one. This behavior results in responses to the follow up question that are more extreme than responses to a certain second cost. To see why, consider subjects whose values for the good lie in between the first and second costs. Under certainty, such a subject will respond either "No-Yes" or "Yes-No" to the two valuation questions, depending on whether the first offered cost is higher or lower than the subject's value. However, in the weighted average case, subjects who respond "No" initially are more likely to respond "No" a second time, since the weighted average of the two costs exceeds the lower bound of the cost range. Similarly, subjects who respond "Yes" initially are more likely to respond "Yes" to the follow up question as well since the weighted average is less than the upper bound of the cost range. This hypothesis predicts that there will be more "Yes-Yes" and "No-No" responses to the two questions than if the second cost were known with certainty.

The Uncertain Cost and Weighted Average hypotheses are observationally distinct and, as theories, are relatively clear. We test these two hypotheses in this paper. This is not true of the two further hypotheses offered in CGM, which we label the Bargaining hypothesis and the Quantity/Quality Shift hypothesis.

The Bargaining hypothesis is based on the intuition that respondents may react to the second stated cost amount by assuming that they are now in a bargaining situation. This is particularly plausible if one remembers that the person implementing the survey is responding as if he or she were bargaining with the respondent (i.e. if the respondent says "Yes"' to the first stated cost amount, the cost amount is raised, but the cost amount is lowered if the respondent says "No").

CGM suggest this intuition implies that respondents are more likely to say "No" to the second cost amount because they hope it will lead to a lowering of any subsequent offers.

CGM do not offer an explicit bargaining model and their stated arguments imply the same predictions as the Uncertain Cost hypothesis. In fact, psychological intuitions aside, there appears to be no difference between these two hypotheses. Consider a two person bargaining situation where neither knows the other's preferences. When considering a response to any offer, the information available to a respondent is the set of possible cost options, a set of beliefs about those options (i.e. a probability distribution over the cost options) and preferences that include attitudes towards the good and towards risk. In other words, the model becomes similar to the model in the Uncertain Cost hypothesis. Thus, our experimental test of the Uncertain Cost hypothesis is also an indirect test of the Bargaining hypothesis, where it is assumed that the subject believes the cost amounts are normally distributed.

The Quantity/Quality Shift hypothesis is based on the intuition that respondents take the price signal to be an indicator of the amount of good on offer or the quality of the good. When offered the second cost amount, the respondent takes this to indicate that a new good is being offered since it has a different price. Thus those who are offered a higher cost amount in the second round are more likely to say "Yes" because they feel that a new, better or bigger, good is being offered. Likewise, those who are offered a lower second cost amount are more likely to say "No" because they feel they are being offered a new, lower quality, or smaller, good.

We find it difficult to interpret this hypothesis from an economic perspective, although as stated, it appears observationally equivalent to the Weighted Average hypothesis. Consider respondents who are offered lower cost amounts, so that they believe that the good now being

offered is of lower quality. *A priori,* their responses depend upon their preferences over this good relative to other goods, their beliefs about the quality (or quantity) change and the amount the price has changed. In fact, if the price has fallen enough, it may *increase* the likelihood of wanting to pay for the good. This is another way of saying that the response to a change in price and quantity of a good depends on the own-price elasticity of demand for that good. Without further information, we feel this hypothesis requires very strong assumptions about the price elasticity of environmental goods in order to make clear predictions about responses. Thus, we do not test this hypothesis in the experiment.

Table I summarizes the predictions of these four hypotheses. There are four possible response combinations to the two valuation questions: "Yes-Yes," "Yes-No," "No-Yes," and "No-No." The first column depicts the number of each type of response predicted under certainty and truthful preference revelation. A, B, C, and D represent non-negative amounts of each response type that depend on the distribution of values for the public good in the sample population. The values in columns 2, 3, 4, and 5 reflect how the observed responses for each hypothesis are predicted to differ from the certainty, truthful preference revelation baseline. Clearly, the Uncertain Cost and Bargaining hypotheses predict similar deviations from the baseline, as do the Weighted Average and Quantity/Quality Shift hypotheses. Columns 1, 2, and 4 are relevant to the description of the experimental design that follows. Because these four hypotheses result in only two observationally distinct vote distributions, we test only two joint hypotheses: Uncertain Cost/Bargaining and Weighted Average/Quantity-Quality Shift. We induce incentive structures consistent with those described in the Uncertain Cost and Weighted Average hypotheses to test these two compound hypotheses.

TABLE I ABOUT HERE

3. EXPERIMENTAL DESIGN

One of the difficulties with replicating the incentives present in a double-bounded referendum in an experimental laboratory is that subjects in the field do not know that they are going to be asked a second valuation question until the moment that the question is posed to them. Implementing this aspect of contingent valuation surveys generates problems in a controlled laboratory setting where, to conform to proper experimental protocol, subjects must be fully informed of experimental procedures from the outset ([8], p. 23). Thus, while from the perspective of a CV practitioner or a behavioral economist, a sequential, double referendum might seem the most natural and appropriate design, it is not only problematic but also incorrect.

It is problematic since telling subjects at the beginning that they will be asked two questions can generate strategic responses to the first question, particularly if subjects know that they will receive a lower cost offer in the second question if they respond "No" to the first question. Since the hypotheses we wish to test require truthful preference revelation in the first question, it is not possible to incorporate this type of two-question approach into the design of the experiment.

One potential solution to the problem of potential strategic responses to the first question would be to design the experiment in two parts. Subjects play part one, which corresponds to the first referendum question, with the knowledge that the experiment has two parts and that each part is played as a separate game. However, they would not know the rules of the game for part two until after they completed part one. Part two would then correspond to the theoretical conditions described by CGM.

This approach is incorrect because it will not always be possible to create the incentive structures described by CGM. Their hypotheses center solely around responses to the second

offer, conditional on truthful responses to the first. A sequential design cannot guarantee that subjects will always respond truthfully to the first question. In fact, some studies (e.g. [13]) that find no difference in behavior in real and hypothetical referenda indicate that this lack of bias is at least in part a result of non-incentive compatible behavior in real referenda such as what would comprise part one of such a sequential experiment. This result is troubling, as the assignment of costs in part two of the experiment must be conditional on the part one response. It is not clear how to assign a cost in part two to a subject who gives a non-incentive compatible response in part one. As a result, a sequential design could result in outcomes different from CGM's predictions simply because subject behavior in the first round deviates from that specified by the theory. While such behavior is certainly an interesting issue, it is not the focus of this experiment. The purpose here is to investigate responses to the second question of a double referendum under the specific conditions of uncertainty described by CGM. Therefore, we require a design that meets the criteria specified in the theoretical models.

Consequently, we employ a "virtual" first round followed by a one-shot referendum. Here, subjects play only one round under the (experimental) assumption that they have played a previous one-shot round in which they truthfully revealed their preferences, consistent with the (econometric) assumption that respondents truthfully reveal their preferences when responding to the first bid amount in a referendum contingent valuation study [5]. Further, [4] provide experimental evidence to support this assumption.

The first round is virtual in the sense that subjects do not actually play it. However, given their assigned value and cost and the assumption of truthful preference revelation, it is possible to predict what subjects would have voted in this round. This "virtual first vote" then determines the cost that subjects are offered in the second, or actual round, which is the only round that

subjects play. For example, if a subject's assigned value exceeds the initial offered cost, his or her virtual first vote is "Yes." The subject is then offered a cost higher than the initial cost in the actual round. If a subject's assigned value is less than the first offered cost, his or her virtual first vote is "No." The subject will then be offered a cost lower than the initial cost in the actual round. Figure 1 depicts how subjects' assigned values, costs, and virtual votes determine their cost assignments in the actual round of the experiment. We describe the details of the experimental design below.

FIGURE 1 ABOUT HERE

3.1. General Experimental Procedures

Two hundred sixteen cadets from the United States Air Force Academy participated in this experiment. The sample size was determined via a power analysis using a range of assumptions about observed deviations between treatments, ranging from small to large. Under very conservative assumptions, a sample size of 200 results in a power of approximately 0.80 at a level of significance of 0.05. At this sample size, the power approaches one rapidly as deviations between treatments increase. Because of the experimental design, the sample size must be divisible by 27. Therefore, the sample size is 216. Table II outlines the design matrix for the experiment.

Subject recruitment was accomplished using posters and flyers and announcements in classes. Subjects were informed of the amount of time the experiment would take (approximately 30-45 minutes) and average earnings ($10.00).[1]

TABLE II ABOUT HERE

At each experimental session, 27 students arrive at a designated room at the designated time. Security measures are in place to ensure that no subject participates more than once. Subjects are then randomly assigned to one of the three treatments, each of which occurs in a different room. By running all three treatments concurrently at each experimental session, we can reduce session-specific effects by randomly distributing subjects among treatments. Once in the different rooms, each subject receives a blank envelope containing an instruction packet and a smaller envelope, which contains a ballot slip and "Investment Return Slip." The information in the experimental materials includes each subject's assigned induced values. These assignments were also randomly drawn from the overall distribution of induced values described below.

The experimental design consists of three treatments: the Baseline treatment, the Uncertain Cost treatment, and the Weighted Average treatment. In each treatment, the basic design of the experiment is the same. Each of the 9 subjects is a shareholder in an imaginary investment company called "Experimental Investments Limited (EIL)." The group must vote on whether EIL should make an investment for the group. Each shareholder begins the experimental session with a total of 10 tokens. Each shareholder knows how his or her cost of the investment will be determined and personal return from the investment. EIL makes the investment if at least 6 of the 9 shareholders vote "Yes." If less than 6 shareholders vote "Yes," then the session terminates and each shareholder completes the experiment with 10 tokens.[2] If the vote passes, then shareholders' final token balance equals 10 minus their investment amount plus their return from the investment. At the completion of a session, tokens are converted to dollars at the rate of $1.00 per token. The only variable that differs between treatments is how shareholders' individual costs of the investment are determined. As a result, the instructions for each treatment

differ in their description of this variable of interest. The Appendix contains the relevant portions of the experimental instructions for each treatment.[3]

Table III presents the return and cost structures for the three treatments. Table IV reports the predicted distributions of "Yes-Yes," "Yes-No," "No-Yes," and "No-No" votes that result from the design of the virtual first round, the induced values and costs in the actual round, and predictions of the hypotheses proposed by CGM. The high and very high value subjects' values and costs determine the predicted numbers of "Yes-Yes" and "Yes-No" votes. The low value subjects' values and costs determine the predicted numbers of "No-Yes" and "No-No" votes. There are no predicted "No-No" votes in the Baseline treatment because we omit this subject type from the experimental design. However, both the Uncertain Cost and Weighted Average hypotheses predict a positive number of "No-No" votes in this treatment. Therefore, nonzero levels of "No-No" votes are one form of evidence for behavior that supports these hypotheses.

TABLE III ABOUT HERE

TABLE IV ABOUT HERE

3.2. Detailed Treatment Procedures

In the Baseline treatment, shareholders know their cost of the investment with certainty. Therefore, shareholders whose return exceeds their assigned cost are expected to vote "Yes" and shareholders whose return is less than their assigned cost are expected to vote "No." This treatment serves as the control and expected outcomes are identified in column 1 of Table IV.

In the Uncertain Cost treatment, shareholders know that their cost of the investment will be drawn from a normal distribution. Although CGM do not explicitly state that respondents' uncertainty about the cost takes the form of a normal distribution, we take as our starting point

the view that any theory that formalizes this uncertainty would be unlikely to exclude the normal distribution. In the experiment, the normally distributed cost has a mean (MeanCost) as described in Table III and a standard deviation of 5. Shareholders know their assigned MeanCost and see a table illustrating how their cost of the investment will be determined based on a drawing from a bag of cards numbered from 1 to 100. This table is contained in the experimental instructions in the Appendix. The drawing occurs after the vote and only if 6 of the 9 shareholders vote "Yes." It is important to note that the possible values of MeanCost equal the assigned costs that are known with certainty in the Baseline treatment. Therefore, this treatment replicates the incentive structure described in the Uncertain Cost hypothesis. Column 2 of Table IV reports the predicted outcomes from this treatment.

In the Weighted Average treatment, shareholders know that their cost of the investment will be either a low value (LowCost) or a high value (HighCost). Shareholders know their values of LowCost, HighCost, and their return from the investment. They also know that their cost will be determined by two drawings from a bag of cards numbered from 1 to 100. The first drawing occurs before the vote and determines the "threshold" number. The threshold number randomly partitions the [1,100] interval into two segments. The first is [1, threshold] and the second is [threshold + 1, 100]. The result of the first drawing sets the stage for a second drawing from the same pool of numbers that occurs only if the vote passes. A draw from the first segment results in a shareholder cost of LowCost, and a draw from the second segment results in a shareholder cost of HighCost. The larger is the first "threshold" draw, the higher is the probability that the second draw will result in a LowCost outcome. Shareholders see a table summarizing the rules of the drawing, which is included in the Appendix. The values of LowCost and HighCost correspond to the offered costs in the first and second valuation questions as depicted in Figure

1. Whether a shareholder's second cost is lower or higher than the first depends on the virtual vote, which in turn depends on the assigned return from the investment. This treatment replicates the incentive structure described in the Weighted Average hypothesis. Predicted outcomes are presented in column 3 of Table IV.

4. RESULTS AND DISCUSSION

Table V reports the observed vote distributions in each treatment. A cursory examination of the table indicates that subjects do behave in accordance with the predictions of economic theory in the case of the Baseline treatment. Furthermore, the vote distribution in the Weighted Average treatment does move in the direction predicted by the hypothesis. However, this is not the case for the Uncertain Cost treatment. In particular, there are zero "No-No" votes in this treatment. While this hypothesis predicts that the distribution will be shifted to the left of that resulting from the Baseline treatment, the data seem to indicate a shift to the right. We can quantify these observations by testing the following hypotheses using Pearson's chi-squared and Fisher's exact tests as described in Agresti [1]:

Hypothesis 1: The vote distributions in the Baseline and Uncertain Cost treatments are the same.

Hypothesis 2: The vote distributions in the Baseline and Weighted Average treatments are the same.

Failure to reject hypotheses 1 and 2 provides evidence that the changes in cost structures across treatments have no effect on subject behavior relative to that observed in the Baseline treatment. Table VI reports the results of tests of hypotheses 1 and 2.[4]

We find weak evidence to reject Hypothesis 1. In fact, we can only reject this hypothesis at the 10% level of significance ($p = 0.072$). There are two explanations for the weak rejection of this hypothesis. The first is that subjects simply do not behave in accordance with the theoretical predictions associated with the Uncertain Cost hypothesis. The second is that one of the assumptions of the theory is violated. In particular, CGM argue that one will only observe an increase in "No" responses to the second question for risk averse subjects. One possible explanation for this result is that this subject pool contains few risk averse subjects.

A simple test for risk aversion conducted on a subsample of USAFA cadets revealed that approximately half of the subjects were risk loving or risk neutral. [5] However, since there are no individual identifiers with the experimental data, it is not known how many, if any, of these subjects participated in the Uncertain Cost treatment. If the risk preferences in the subsample that participated in the risk aversion test are representative of those in the cadet population, then the results of this test provide some insight into the outcome of the Uncertain Cost treatment. Clearly, the risk preferences of respondents are an important variable to consider if uncertainty is an issue in a survey. Knowledge of these preferences would be crucial to proper analysis of the data. In summary, there is weak evidence to suggest that subjects invoke the Uncertain Cost hypothesis, albeit the direction of the effect is the opposite of that proposed by CGM. This result indicates that further investigation of the hypothesis in a setting where the risk preferences of subjects are known is warranted. [6]

We find strong evidence to reject Hypothesis 2. We can reject the null hypothesis that the vote distributions in the Baseline and Weighted Average Treatments are the same at any standard level of significance ($p = 0.003$). We further note that the distribution is as predicted by CGM. Therefore, subjects do behave in accordance with the predictions of the Weighted Average

hypothesis. If subjects take a weighted average of the presented costs when responding to a double-bounded field survey, then one should expect to obtain estimated willingness to pay distributions that are consistent with our results.

In the presence of cost averaging, econometric methods can be applied to explicitly incorporate this information about subject behavior into the estimation procedure.[7] It is interesting to note that such a model has similarities to one developed by Herriges and Shogren [11] to describe a model of anchoring in double referenda in which subjects update their WTP for the good based on the value of the first bid amount. If subjects behave in this manner in response to double referenda, their behavior effectively widens the bounds placed on willingness to pay by the follow-up question. By contrast, if subjects' behavior is consistent with the weighted averaging model, the result is exactly the opposite. Subject responses to the follow-up question effectively narrow the bounds on willingness to pay.

TABLE V ABOUT HERE

TABLE VI ABOUT HERE

5. CONCLUSIONS

The experiment reported above was not designed to prove conclusively which, if any, of the tested hypotheses are adopted by individual respondents when faced with a follow up bid in a double referendum. What we have done is to test the two proposed hypotheses under strict laboratory conditions to examine whether, when strongly induced, they are capable of generating the types of voting behavior both expected in theory and observed in the contingent valuation literature. Our experimental findings, while not testing every possible explanation of apparently inconsistent observed responses, are a useful starting point in providing much needed empirical evidence in the area.

The implications of the results from the Weighted Average treatment are straightforward. If such behavior is present within a particular sample of contingent valuation survey respondents then there will be significant differences in voting behavior over what would be the case compared to certain costs (i.e. our Baseline treatment). This difference is in the direction proposed by CGM. It is, however, difficult to draw definitive conclusions with respect to the Uncertain Cost hypothesis. As currently formulated, CGM's theory predicts that for risk averse subjects the difference in voting behavior between the Baseline and Uncertain Cost treatments will result in more "No" responses to the second question, as illustrated in Table I. Although we did observe significant differences at the 10% level between treatments, they were in the opposite direction to those proposed by CGM. A sample that included a sufficient number of risk loving respondents would produce such a result. To date, as far as we are aware, there is no evidence regarding the risk preferences of past contingent valuation survey respondents. Therefore, it is difficult to say whether the Uncertain Cost hypothesis has had any impact on contingent valuation survey results, whether in the direction predicted by CGM or in the direction found in our study.

In closing we note that this paper demonstrates how appropriate procedures from mainstream experimental economics can be utilized to test previously unverified explanations of econometrically observed inconsistent voting in double referenda contingent valuation surveys. If subject behavior matches the theoretical predictions, and if the same outcomes are observed in the field, then these experimental results provide information for use in developing appropriate econometric techniques to apply to data from double bounded referendum surveys. For example,

if field data exhibit response patterns consistent with the Weighted Average hypothesis (more "Yes-Yes" and "No-No" responses than predicted), then it may be reasonable to incorporate information about the weighted average of the two offered costs into the log likelihood function used in the estimation procedure.

In addition, these results may be ultimately helpful in redesigning double referendum surveys to avoid the respondent uncertainty and resulting bias typically associated with the follow up bid. While the results of this study are unable to generate specific recommendations in this regard, they do point out the direction for further research. In order to generate information that will be useful in redesigning surveys, an experimental research agenda must answer three questions:

(1) Is subject behavior consistent with the theoretical predictions when the conditions described in the theoretical model are induced? (Theory falsification)

(2) Which theories seem to correspond to subject behavior in the field? (Theory selection)

(3) How do alternative survey designs impact subject behavior? (Survey modification)

The purpose of these experiments has been to begin to address question (1). Future experimental research agendas may address alternative theories of behavior in double referenda, as well as address questions (2) and (3). An alternative to question (3) is the development of new econometric estimators, such as the one described above, to model the behaviors identified in question (2). One response to question (3) may be to modify survey designs such as the double referendum to reduce possible impacts of design upon subject behavior. Another response described in this paper would be to develop new econometric estimators which model the actual subject behavior in double referenda as identified in question (2). Ideally, these two lines of

research would occur in parallel. A systematic experimental research agenda such as the one outlined above can result in useful information for survey practitioners.

REFERENCES

1. A. Agresti, *Categorical Data Analysis*, John Wiley & Sons, New York (1990).

2. A. Alberini, Efficiency vs. bias of willingness-to-pay estimates: Bivariate and interval-data models, *J. Environ. Econom. Management* **29**(2), 169-180 (1995), doi:10.1006/jeem.1995.1039.

3. K.J. Arrow, R. Solow, P. Portney, E. Leamer, R. Radner, and H. Schuman, Report of the NOAA panel on contingent valuation, *Federal Register* **58(**10), 4602-4614 (January 15, 1993).

4. A.C. Burton, K.S. Carson, S.M. Chilton, and W.G. Hutchinson, Consequentiality and the incentive compatibility of one-shot referenda, Newcastle Discussion Papers in Economics No. 2001-01.

5. T.A. Cameron, and J. Quiggin, Estimation using contingent valuation data from a 'dichotomous choice with follow-up' questionnaire, *J. Environ. Econom. Management* **27**(3), 218-234 (1994), doi:10.1006/jeem.1994.1035.

6. R. Carson, T. Groves, and M. Machina, Incentive and informational properties of preference questions, Plenary Address to the European Association of Resource and Environmental Economists, Oslo, Norway, June 1999.

7. R.G. Cummings, S. Elliott, G.W. Harrison, and J. Murphy, Are hypothetical referenda incentive compatible? *J. Political Econom.* **105**(3), 609-621 (1997).

8. D.D. Davis and C.A. Holt, "Experimental Economics," Princeton University Press, Princeton, NJ (1993).

9. J.R. DeShazo, Designing transactions without framing effects in iterative question formats, *J. Environ. Econom. Management*, 43(3), 360-385 (2002), doi:10.1006/jeem.2000.1185.

10. M. Hanemann, J. Loomis, and B. Kanninen, Statistical efficiency of double-bounded dichotomous choice contingent valuation, *Amer. J. Agr. Econom.* **73**(4),1255-1263 (1991).

11. J.A. Herriges and J.F. Shogren, Starting point bias in dichotomous choice valuation with follow-up questioning, *J. Environ. Econom. Management.* **30**(1), 112-131 (1996), doi:10.1006/jeem.1996.0008.

12. D. McFadden, and G.K. Leonard, Issues in the contingent valuation of environmental goods: Methodologies for data collection and analysis, in "Contingent Valuation: A Critical Assessment" (J. Hausman, Ed.), North-Holland, Amsterdam (1993).

13. L.O Taylor, M. McKee, S.K. Laury, and R.G. Cummings, Induced value tests of the referendum voting mechanism, *Econom. Letters* **71**(1), 61-65 (2001).

FOOTNOTES

[1] Weekly cadet take-home pay ranges from $65-$100. Therefore, an additional $10 for less than an hour's time represents a significant addition to a cadet's weekly income. Given this, there is no doubt that cadets took the experiment seriously.

[2] Students were made aware in the introduction to the experiment that while the experimental currency was tokens, these would be converted into dollar amounts at the end of the session.

[3] A complete package of all experimental materials is available from the authors upon request.

[4] It is useful to note that the vote distributions in the Weighted Average and Uncertain Cost treatment are also significantly different from each other at the 10% level ($p = 0.071$)

[5] Subjects make a series of ten choice decisions between two gambles, A and B. Gamble A has a certain payoff of $3. Gamble B has a payoff of either $6 or $1. The probability of the $6 payoff changes from 0.1 to 1 across the ten choice decisions. After subjects make the ten choices, a ten-sided die is rolled twice. The first roll determines which of the ten choices will be used to compute the subjects' earnings from the experiments. The second roll determines whether subjects who chose gamble B for the relevant choice decision earn $6 or $1. Subjects' risk preferences can be determined by examining at what probability the subjects switch from choosing the certain gamble to choosing the risky one. Instructions for the risk experiment are also available from the authors upon request.

[6] We attempted to re-run the Uncertain Cost treatment with a group of subjects at Queen's University, Belfast, who were pre-screened to be risk averse. However, only 8 of the 28 subjects who showed up for the experiment were determined to be risk averse. It may very well be the case that those who self-select to participate in economics experiments may be more risk loving.

[7] Given the following:

$WTP = f(X;\theta,\varepsilon) = g(X;\theta,\varepsilon) = exp(\alpha)\varepsilon$, where ε follows a lognormal distribution.

B_{1i} = the first bid amount offered to subject i

B_{Li} = the second bid amount offered to subjects who respond "No" to B_{1i}, $B_{Li} < B_{1i}$.

B_{Hi} = the second bid amount offered to subjects who respond "Yes" to B_{1i}, $B_{Hi} > B_{1i}$.

γ = the weight that subjects place on the first bid. $\gamma \in [0,1]$.

I_{1i} = 1 if subject i responds "No, No" to the two WTP questions, 0 otherwise.

I_{2i} = 1 if subject i responds "No, Yes" to the two WTP questions, 0 otherwise.

I_{3i} = 1 if subject i responds "Yes, No" to the two WTP questions, 0 otherwise.

I_{4i} = 1 if subject i responds "Yes, Yes" to the two WTP questions, 0 otherwise.

The log-likelihood function for responses to double referenda with weighted averaging behavior is:

$$\ell = \sum_{i=1}^{n} I_{1i} \ln\left\{\Phi \frac{[\gamma \ln B_{1i} + (1-\gamma)\ln B_{Li} - \alpha]}{\sigma}\right\} + I_{2i} \ln\left\{\Phi\left[\frac{\ln B_{1i} - \alpha}{\sigma}\right] - \Phi\left[\frac{\gamma \ln B_{1i} + (1-\gamma)\ln B_{Li} - \alpha}{\sigma}\right]\right\}$$
$$+ I_{3i} \ln\left\{\Phi\left[\frac{\gamma \ln B_{1i} + (1-\gamma)B_{Hi} - \alpha}{\sigma}\right] - \Phi\left[\frac{\ln B_{1i} - \alpha}{\sigma}\right]\right\} + I_{4i} \ln\left\{1 - \Phi\left[\frac{\gamma \ln B_{1i} + (1-\gamma)\ln B_{Hi} - \alpha}{\sigma}\right]\right\}$$

A complete derivation of this function is available upon request from the authors.

TABLE I

Double Referendum Response Effects

	Certainty, Truthful Preference Revelation (Baseline)	***Uncertain Cost***	***Bargaining***	***Weighted Averaging***	***Quantity/ Quality Shift***
Yes-Yes	A	< A	< A	> A	> A
Yes-No	B	> B	> B	< B	< B
No-Yes	C	< C	< C	< C	< C
No-No	D	> D	> D	> D	> D

TABLE II

Experimental Design Matrix

	Baseline	**Weighted Average**	**Uncertain Cost**	**Totals**
Treatments per Session	1	1	1	3
Subjects per Session	9	9	9	27
x *Number of Experimental Sessions*	x 8	x 8	x 8	x 8
= Total Subjects	72	72	72	**216**

TABLE III

Assignment of Induced Values by Treatment

	Baseline	**Weighted Average**	**Uncertain Cost**
Low Value Shareholder Type			
Return	16 tokens	16 tokens	16 tokens
Cost Structure	Certain cost of 10 tokens	Cost is either LowCost or HighCost Determined by Drawing LowCost = 10 tokens HighCost = 20 tokens	Cost is normally distributed Mean = MeanCost, $\sigma = 5$ Determined by Drawing MeanCost = 10 tokens
Subjects of this type per Treatment	24	24	24
High Value Shareholder Type			
Return	26 tokens	26 tokens	26 tokens
Cost Structure	Certain cost of 30 tokens	Cost is either LowCost or HighCost Determined by Drawing LowCost = 20 tokens HighCost = 30 tokens	Cost is normally distributed Mean = MeanCost, $\sigma = 5$ Determined by Drawing MeanCost = 30 tokens
Subjects of this type per Treatment	24	24	24
Very High Value Shareholder Type			
Return	32 tokens	32 tokens	32 tokens
Cost Structure	Certain cost of 30 tokens	Cost is either LowCost or HighCost Determined by Drawing LowCost = 20 tokens HighCost = 30 tokens	Cost is normally distributed Mean = MeanCost, $\sigma = 5$ Determined by Drawing MeanCost = 30 tokens
Subjects of this type per Treatment	24	24	24

TABLE IV

Predicted Responses by Treatment

	Baseline	**Uncertain Cost**	**Weighted Average**
Vote Type			
Yes-Yes	24	< 24	> 24
Yes-No	24	> 24	< 24
No–Yes	24	< 24	< 24
No-No	0	> 0	> 0

TABLE V

Observed Responses by Treatment

	Baseline	**Uncertain Cost**	**Weighted Average**
Vote Type			
Yes-Yes	24	35	36
Yes-No	24	13	12
No–Yes	24	24	18
No-No	0	0	6

TABLE VI

Hypothesis Test Results

Hypothesis 1 – H_0: Baseline = Uncertain Cost	
Pearson's Chi-Squared: p = 0.070	Fisher's Exact: p = 0.072
Hypothesis 2 – H_0: Baseline = Weighted Average	
Pearson's Chi-Squared: p = 0.004	Fisher's Exact: p = 0.003

APPENDIX

Experimental Instructions Excerpts

1. Baseline Treatment Instructions
2. Uncertain Cost Treatment Instructions
3. Weighted Average Treatment Instructions

Experimental Investments Limited and Its Procedures
Each of you is a shareholder in Experimental Investments Limited (EIL). As a shareholder, you will have the opportunity to vote on whether EIL will invest a number of tokens on your behalf. To help you understand the investment we will use the following abbreviations:

C = your investment amount in tokens

R = your return from the investment in tokens

If EIL votes to make the investment, then **C** tokens will be deducted from your account balance. You will know your value of **C** before you vote on the investment decision. Different people in the room have different values of **C**. This value of **C** is private information. Do not reveal your value of **C** to the other people in the room.

If EIL votes to make the investment, then **R** tokens will be added to your account balance. You will know your value of **R** before you make the investment decision. Different people in the room have different values of **R.** This value of **R** is also private information. Do not reveal your value of **R** to the other people in the room.

To summarize, if EIL votes to make the investment then **C** tokens will be deducted from your account balance and **R** tokens will be added to your account balance. When you make your investment decision you will know your values of **C** and **R**.

Are there any questions?

Please turn the page.

Experimental Investments Limited and Its Procedures (cont.)
The EIL procedure is:

(i) Each of you will open Envelope 1, which contains your VOTING SLIP. You will see your values of **C** and **R** on your VOTING SLIP. DO NOT show these values to the other individuals in the room.

(ii) Each of you will fill in the VOTING SLIP to indicate whether or not you wish EIL to invest **C** tokens on your behalf. Voting will be in private so the other members of EIL will not know how you voted.

(iii) I will collect everyone's VOTING SLIP in the ballot box and count the votes. You will be able to verify this count at the end of the experimental session if you so desire.

If less than 6 of you vote 'YES' to the proposal for EIL to invest **C** tokens on your behalf, then EIL will not make the investment. No deductions will be made from anyone's account balance and no returns will be added to anyone's account balance.

If 6 or more of you vote 'YES' to the proposal for EIL to invest **C** tokens on your behalf, then EIL will make the investment. As a shareholder, **C** tokens will be deducted from your account balance and **R** tokens, your return from the investment, will be added to your account balance. Even if you voted 'NO,' EIL will still invest **C** tokens on your behalf and you will still receive **R** tokens.

Are there any questions?

Please turn the page.

Experimental Investments Limited and Its Procedures
Each of you is a shareholder in Experimental Investments Limited (EIL). As a shareholder, you will have the opportunity to vote on whether EIL will invest a number of tokens on your behalf. To help you understand the investment we will use the following abbreviations:

C = your investment amount in tokens

M = the average value of **C** for you

R = your return from the investment in tokens

If EIL votes to make the investment, then **C** tokens will be deducted from your account balance. The value of **C** depends upon your value of **M** and the results of a drawing that I will conduct if EIL votes to make the investment. I will explain how the drawing will work on the next page.

You will know your value of **M** before you vote on the investment decision. Different people in the room have different values of **M**. This value of **M** is private information. Do not reveal your value of **M** to the other people in the room.

If EIL votes to make the investment, then **R** tokens will be added to your account balance. You will know your value of **R** before you vote on the investment decision. Different people in the room have different values of **R**. This value of **R** is also private information. Do not reveal your value of **R** to the other people in the room.

Are there any questions?

Please turn the page.

The Drawing
The outcome of a drawing will determine your investment amount. We will only conduct this drawing if EIL votes to make the investment decision.

The bag that I am holding contains 100 cards numbered from 1 to 100. You may verify the numbers on the cards at any time during the experiment if you so desire. If EIL votes to make the investment, I will ask one of you to draw one card from this bag and announce the number on the card. This number will determine your investment amount, **C**.

The table below explains how the value of **C**, your investment amount, depends upon the outcome of the drawing. For example, if the number on the card drawn is 33, then your investment amount is **M-2**. If the number on the card drawn is 86, then your investment amount is **M + 5**.

You might find it helpful to know that the numbers in this table are based on a normal distribution with a mean of **M** and a standard deviation of 5.

If the card number is:	**Then the value of C is:**
1	**M -10**
2 or 3	**M - 9**
4 or 5	**M - 8**
6, 7, or 8	**M – 7**
9, 10, 11, or 12	**M - 6**
13, 14, 15, 16 or 17	**M - 5**
18, 19, 20, 21, 22, or 23	**M - 4**
24, 25, 26, 27, 28, 29, or 30	**M - 3**
31, 32, 33, 34, 35, 36, 37, or 38	**M - 2**
39, 40, 41, 42, 43, 44, 45, or 46	**M - 1**
47, 48, 49, 50, 51, 52, 53, or 54	**M**
55, 56, 57, 58, 59, 60, 61, or 62	**M + 1**
63, 64, 65, 66, 67, 68, 69, or 70	**M + 2**
71, 72, 73, 74, 75, 76, or 77	**M + 3**
78, 79, 80, 81, 82, or 83	**M + 4**
84, 85, 86, 87, or 88	**M + 5**
89, 90, 91,or 92	**M + 6**
93, 94, or 95	**M + 7**
96 or 97	**M + 8**
98 or 99	**M + 9**
100	**M + 10**

Please turn the page.

Experimental Investments Limited and Its Procedures
Each of you is a shareholder in Experimental Investments Limited (EIL). As a shareholder, you will have the opportunity to vote on whether EIL will invest a number of tokens on your behalf. To help you understand the investment we will use the following abbreviations:

L = your low investment amount in tokens

H = your high investment amount in tokens

R = your return from the investment in tokens

If EIL votes to make the investment, then either **L** or **H** tokens will be deducted from your account balance. The outcomes of two drawings will determine whether the investment amount deducted from your account is **L** or **H**. I will explain how the drawings will work on the next page.

You will know your values of **L** and **H** before you vote on the investment decision. Different people in the room have different values of **L** and **H**. These values of **L** and **H** are private information. Do not reveal your values of **L** and **H** to the other people in the room.

If EIL votes to make the investment, then **R** tokens will be added to your account balance. You will know your value of **R** before you vote on the investment decision. Different people in the room have different values of **R**. This value of **R** is also private information. Do not reveal your value of **R** to the other people in the room.

Are there any questions?

Please turn the page.

The Drawings
The outcomes of two drawings will determine whether your investment amount is **L** or **H**. The purpose of the first drawing is to determine the *threshold number*. The *threshold number* indicates whether it is more likely that your investment amount will be **L** or that your investment amount will be **H**. We will conduct this drawing before you vote to make the investment decision.

The purpose of the second drawing is to determine which amount, **L** or **H**, is your investment amount. We will only conduct the second drawing if EIL votes to make the investment decision.

The bag that I am holding contains 100 cards numbered from 1 to 100. You may verify the numbers on the cards at any time during the experiment if you so desire. Before EIL votes on the investment decision, I will ask one of you to draw one card from the bag and announce the number on the card. This number will determine the *threshold number*.

The table below explains how the *threshold number* indicates the likelihood that your investment amount will be **L** or **H**. If EIL votes to make the investment, we will conduct a second drawing to determine whether your investment amount is **L** or **H** in the same manner as we conduct the first drawing.

If the card that we draw in the second drawing is numbered from 1 to the *threshold number*, then **L** tokens will be deducted from your account balance.

If the card that we draw in the second drawing is numbered from the *threshold number plus one* to 100, then **H** tokens will be deducted from your account balance.

You might find it helpful to know that the likelihood that your investment amount is **L** or **H** is based on a weighted average of **L** and **H**.

If the card that we draw in the second drawing is numbered	***Then your investment amount is:***
From 1 to the *threshold number*.	**L**
From the *threshold number plus one* to 100	**H**

Please turn the page.