Testing for Budget Constraint Effects in a National Advisory Referendum Survey on the Kyoto Protocol

Hui Li, Robert P. Berrens, Alok K. Bohara, Hank C. Jenkins-Smith, Carol L. Silva, and David L. Weimer

In contrast to providing standard reminders about remembering household budgets, does asking survey respondents about their discretionary income and its use affect their voting responses in a national advisory referendum survey? We explore this question using U.S. household data from a unique set of multi-mode random samples (telephone and Internet surveys), and an advisory referendum concerning the Kyoto Protocol. The contingent valuation method is applied to estimate household willingness to pay (WTP) for a split-sample treatment: respondents who only received a standard reminder of household budgets (control group) versus respondents who received two mental accounting-type questions on discretionary income and its uses (treatment group). Results indicate that the treatment significantly influences voting responses and lowers estimated household WTP.

Key words: budget constraint, contingent valuation, Kyoto Protocol, mental accounts, referendum

Introduction

Given critical information gaps in a variety of cost-benefit analyses, natural resource damage and other planning assessments, the demand for nonmarket values for changes in environmental goods persists (Bishop, 2003; List et al., 2004). But concerns also persist that stated preference approaches, such as the survey-based contingent valuation (CV) method, may be subject to hypothetical bias, especially in the case of complex environmental goods or policy changes (e.g., concerning climate change or biodiversity). While there are various strategies for improving the validity of CV estimates, such as advocating the use of referendum formats, and controlling for response uncertainty, there remains much we don't know and considerable room for improved understanding of preferences for complex environmental goods.[1]

[1] For a meta-analysis of differences between actual and hypothetical stated values, see Little and Berrens (2004).

This empirical investigation uses a national (U.S.) advisory referendum CV format, in which the good being valued is the reduction of global warming through the ratification of the Kyoto Protocol. The Kyoto Protocol is a complex international treaty which, to take effect, required ratification by at least 55 developed countries and at least a 55% reduction in the total 1990 carbon dioxide emissions.[2] The Protocol imposes binding emission reductions in greenhouse gases for developed countries, with no concomitant obligations for developing countries (e.g., no restrictions on increases in emissions for rapidly growing economies such as India and China). The Protocol requires the participating industrialized countries to reduce their greenhouse gas emissions below 1990 levels based on their varying targets during the period from 2008 to 2012.

With implementation of the Protocol beginning in early 2005, significant questions about its effectiveness remain. As the largest industrialized economy in the world, particular concern surrounds U.S. withdrawal from the treaty. The issue of bringing the United States back into serious international negotiations (on a revised Kyoto Protocol or some variant thereof) remains important for future advancements in international coordination on global climate change.[3] Thus, the issue of accurately assessing the degree of U.S. public support for the Kyoto Protocol remains a relevant informational input.

Given this context, this study investigates the issue of possible budget constraint effects in referendum CV responses. Within the broader CV validity debate, there has been considerable concern that respondents may not give full consideration to their ability to pay (e.g., Arrow et al., 1993). Conversely, the credibility of results from a valuation study would be enhanced if there were evidence indicating respondents had taken into account "their personal circumstances and constraints" (Bateman and Langford, 1997, p. 1215). Failure to consider budget constraints is viewed as a source of possible upward bias (Kemp and Maxwell, 1993; Willis and Garrod, 1993). To reduce possible "budget constraint bias," Mitchell and Carson (1989) suggest the use of standard reminders to encourage respondents to fully consider their incomes. Mixed evidence has accumulated concerning the effect of explicit budget constraint (and substitute goods) reminders. Also at issue is what the relevant income is in different circumstances, and whether "mental accounts" explain responses (Bateman and Langford, 1997).

To further our understanding of possible budget constraint effects, the objective of this study is to investigate the effect on referendum voting of having respondents answer two mental accounting-type questions. A unique aspect of the study is the inclusion of both national (U.S.) telephone and Internet samples using a web-based survey. In addition to providing all respondents with a standard budget reminder as part of the elicitation question, we exposed a split-sample treatment group to two questions asking them to consider: (*a*) the proportion of monthly income available for optional uses after main expenses, such as housing; and (*b*) the proportion of their monthly income available for

---

[2] Following a long series of extended international negotiations, as of February 16, 2005, the Kyoto Protocol was to take effect on its 128 parties. (For the announcement, refer to the United Nations website.) Only four industrial countries (out of 59) did not ratify the Kyoto Protocol, but these four (including the United States) account for one-third of total greenhouse gas emissions globally. Thus, it is argued by a number of sources that the absence of U.S. participation, which withdrew its support in 1998, may severely limit the overall effectiveness of the treaty (for further discussion, see Li et al., 2004).

[3] Congressional debate on the U.S. stance on requiring emission reductions for greenhouse gases remains open. As a continuing effort to raise bipartisan concern, Senators John McCain (R) and Joe Lieberman (D) have recently reintroduced the Climate Stewardship Act, with modest U.S. emission reduction targets, for vote in the U.S. Senate. However, the original bill lost 55 to 43 in 2003, and it is expected to face an uphill struggle in 2005 as well (Landler, 2005).

contributions to environmental causes. These questions encourage respondents to construct mental accounts, first in terms of discretionary income for optional uses and then in terms of the use of this discretionary income for environmental causes. Results indicate that, relative to the standard reminder control group, the mental accounts treatment significantly affects voting responses and lowers estimated household willingness to pay (WTP).

Background and Literature Review

The issue of meaningful consideration of budget constraints gained some topical importance in CV research when it was emphasized by the blue-ribbon panel established by the National Oceanic and Atmospheric Administration (NOAA) (Arrow et al., 1993). In the Panel's critique and review of the CV method for estimating passive use values for inclusion in natural resource damage assessments, a number of problems associated with CV studies were noted to be of particular concern. These included the common failure to provide adequate reminders to survey respondents of available substitutes and budget constraints. For WTP elicitation questions, the Panel recommended explicitly: "the payment vehicle must be described fully and clearly with the relevant budget constraint emphasized" (Arrow et al., 1993, p. 4610).

Subsequent to the NOAA Panel report, a number of studies have investigated the effect of explicit reminders of budgets and substitutes on CV valuation responses. Loomis, Gonzalez-Caban, and Gregory (1994) found no significant effect on WTP responses for protecting old-growth forest. They used a split-sample design for the reminder treatment, and a dichotomous choice (DC-CV) format. Adopting a similar format for valuing wetlands preservation, Whitehead and Blomquist (1995) found that information on a set of possible substitutes (complements) had a significant negative (positive) effect on WTP responses. They argued that information on substitutes and complements should be included in any CV study, especially those involving lesser-known goods.

Kotchen and Reiling (1999) attempted to replicate the Loomis, Gonzalez-Caban, and Gregory (1994) study for "lesser-known" public goods of preserving Peregrine falcons and short-nose sturgeons. They used a referendum DC-CV format and an experimental sample design, stratified by the species protected, the information treatment for the substitute, and the budget constraint reminder. The authors found that information on substitutes and budget constraints had no significant effect on mean WTP but did improve estimation efficiency. Whitehead and Blomquist (1999) challenged this conclusion by questioning the statistical performance of the Kotchen and Reiling models (e.g., specification and absence of a significant bid effect in one case). Hailu, Adamowicz, and Boxall (2000) conducted a multi-program CV study on ecosystem conservation. To provide a direct reminder of budget constraints, they asked the respondents to calculate various costs of different combinations of programs. While they found a strong complementary effect among environmental programs, whether respondents considered their *own* budget constraints remains ambiguous.

An alternative to simply providing a split-sample informational treatment in the form of a reminder is to more actively invoke "meaningful consideration" of a budget constraint. Kemp and Maxwell (1993) explored a budget context for CV responses through a "top-down disaggregation" approach. Through mall-intercept surveys, they first asked respondents to answer an open-ended (OE) question for the top-level category of

"national social concerns." Next, respondents were asked a sequence of percentage allocations across different groups of concerns, and for five layers of disaggregation. WTP for the category of interest was then identified (e.g., minimizing risk of oil spills). Kemp and Maxwell found that their top-down disaggregation approach produced distinctly lower WTP for a good when compared to a single-focus approach, which elicits an OE-WTP response solely for the same good. An issue with the results of their study, and their approach, is that by the nature of their design, there are actually multiple public goods changing at once (Smith, 1993). The question is how much of this difference in WTP responses might be due to differences in the two formats in forcing respondents to meaningfully consider budget constraints.

One approach for examining budget constraint effects is provided by Bateman and Langford (1997) in their OE elicitation format CV study of a recreation site. Prior to any WTP questions, respondents were asked to calculate their annual expenditures on all recreational goods; the authors argued that this represented the appropriate "mental account" for the given recreational good. Importantly, Bateman and Langford used a split-sample design, in which the treatment for the budget constraint question was crossed with a number of other treatments. In the various crossed treatments, results always indicated a strong significant positive impact on WTP responses from the presence of the recreational budget constraint treatment. Given significant sample mean values for household recreation budgets, the authors considered this to be a plausible result. The process of getting respondents to consider a response to the relevant mental account is viewed preferable to simple consideration of the household budget, but a priori will have an uncertain impact on WTP responses.

Apart from any consideration of CV surveys, this theory suggests people tend to simplify spending decisions by subdividing a larger budget into various "mental accounts." Both the sources of funds and the uses of funds are labeled in a mental accounting system, and individuals may keep fairly rigid barriers between different accounts (Thaler, 1985, 1990). Thus, in a mental accounting model, expenditures are grouped into categories, such as for housing, food, etc., and then spending is sometimes constrained by these implicit or explicit budgets (Thaler, 1999). The mental accounting model has been offered as an explanation for a number of apparent anomalies in observed consumer choice (Thaler, 1990, 1999). The structure of a mental accounting system may cause an individual to frame a decision in a particular way or from a particular reference point, and this can be a source of loss aversion (Magnussen, 1992). Also, the structure of a mental accounting system may violate the principle of fungibility, causing individuals to focus on relative savings (risk) rather than absolute savings (risk) (Duborg, Jones-Lee, and Loomes, 1997; Moon, Keasey, and Duxbury, 1999).

A number of investigations have considered the effect of mental accounting on CV responses. Specifically, there have been concerns that mental accounting may be the source of a "mental account bias" (Hoevenagel, 1992), also labeled as a type of amenity misspecification, part-whole bias, or a source of scope insensitivity (Magnussen, 1992; Baron, 1997; Duborg, Jones-Lee, and Loomes, 1997). However, such accumulated evidence would not rule out that budget constraint questions help the respondent frame a more reasoned CV response, or that a mental accounting framework appropriately describes the choice process (Bateman and Willis, 1999).

Mental accounting only becomes linked to possible amenity misspecification when: (a) the respondent is employing a mental accounting system, and (b) there is some misunderstanding or miscommunication in the structured CV conversation between surveyor(s) and respondents. Thus, it is appropriate to keep the idea of mental accounting conceptually distinct from amenity misspecification (Magnussen, 1992).

Further, the general notion of mental accounting need not be inconsistent with a model of consumer choice. For example, Magnussen (1992) and Bateman and Langford (1997) note that mental accounting is reminiscent of the idea of two-stage (or n-stage) budgeting. To make this connection more explicit, it is helpful to appeal to consumer theory and the case of implicit separability and two-stage budgeting. The notion of separability requires goods to be partitioned in an individual's utility function, such that preferences within a group can be described independently of the level of goods in other groups (i.e., through a sub-utility function). While conceptually distinct, separability is closely related to the notion of two-stage budgeting. In the latter, individuals are viewed as allocating their income in a hierarchical decision process, such as a two-stage process where budget proportions are first allocated to broad commodity groups (food, entertainment, environmental causes) and then to individual commodities (Deaton and Muellbauer, 1980). So in the environmental valuation literature, two-stage budgeting is consistent with the general notion of mental accounts (Magnussen, 1992), where individuals might be viewed as having a broad expenditure group budget for outdoor recreation, or environmental public goods (Bateman and Langford, 1997). The implication is that CV surveys could include direct questions about discretionary income or budget constraints for broad expenditure groups, as done by Bateman and Langford.

This discussion motivates the experimental design and testing strategy for budget constraint effects pursued in this study. Broadly construed, implementing the notion of mental accounts for expenditure categories or discretionary purposes by actively pursuing this information in a CV survey might encourage respondents to further consider meaningful budget constraints in giving WTP responses. Such questions must be treated as a split-sample informational treatment whose effect can be tested on WTP responses.

The Data and Experimental Design

In order to implement our testing strategy, we use a data set including matching telephone and Internet survey samples (Berrens et al., 2003, 2004). The first sample was collected by a telephone survey (TEL) in January 2000, which was conducted by the Survey Research Center of the Institute for Public Policy at the University of New Mexico. This survey used a random digital dialing (RDD) approach; 1,699 respondents completed the survey, yielding a response rate of 45.6%. In November 2000, a probability-based, Internet survey sample was conducted by Knowledge Networks (KN).[4] KN uses Web TV technology given to the respondents recruited through RDD for

---

[4] The complete data set includes two other large national Internet samples using the Harris Interactive pre-recruited panel, which is non-probability based. For full details of these multi-mode survey samples, see Berrens et al. (2003). The experimental design included two other split-sample treatments not used here, and analyzed elsewhere (see Berrens et al., 2004; Li et al., 2004).

this Internet sample of data; 2,162 respondents completed the survey, yielding a 24.1% response rate.[5]

We focus on a mental accounts treatment, with both the TEL and KN samples. The absence or presence of the mental account treatment is identified by an indicator variable, *MA*. Specifically, a split-sample design was used to randomly divide the respondents into those who received the mental accounts treatment (*MA* = 1) and those who did not (*MA* = 0). Respondents under the mental accounts treatment received two questions in the following order:

- QA: *Now think about your average monthly income and expenses. After you have paid all the necessary bills for such things as housing, transportation, groceries, insurance, debt, and taxes, what percent of your income is left over for optional uses on things like recreation, savings, and giving for charity and other causes?*

- QB: *Now think about the portion of your total income available for optional uses. On average, what percent of that amount do you use for contributions to environmental causes, such as donations for specific programs or contributions and memberships to environmental advocacy groups?*

Thus, respondents under this treatment are encouraged to construct their mental accounting system into income paid for necessary bills (main expenses) and income left over for optional uses. Considering income left for optional uses leads to the first-stage budgeting (QA), while taking into account the income available for contributions to environmental causes leads to the second-stage budgeting (QB). Descriptive statistics for the budget questions are provided in table 1, broken out by the samples for the two survey modes—telephone and Internet (Knowledge Networks).

With respect to the valuation portion of the survey, all respondents were provided with a standard reminder of budget constraints, within the context of the valuation question itself. Specifically, the valuation question uses an advisory referendum format:

- *Suppose that a national vote or referendum were held today in which U.S. residents could vote to advise their Senators whether to support or oppose ratifying the Kyoto Protocol. If U.S. compliance with the treaty would cost your household t dollars per year in increased energy and gasoline prices, would you vote for or against having your Senators support ratification of the Kyoto Protocol? Keep in mind that t dollars spent on increased energy and gasoline prices could not be spent on other things, such as other household expenses, charities, groceries, or car payments.*

For ❑   Against ❑   Don't Know/No Answer ❑

---

[5] Given concerns about potential bias due to the relatively low response rate, we also estimated a full set of matched KN models with weights. Results are consistent with raw data estimations, and are available from the lead author upon request. For the KN sample, a "raking" weighting procedure (conducted by Knowledge Networks) was used based on known demographic marginals for 32 variables, and corrections for any sample selection bias. A comparison of raw and weighted data for the KN sample (and all samples) in the overall study is provided in Berrens et al. (2003). Further, in a mixed-mode data set comparable to ours, Krosnick and Chang (2001) statistically compare a national RDD telephone sample, Harris Interactive sample, and KN sample. They conclude, "Internet-based data collection represents a viable approach to conducting representative sample surveys" (p. 8), especially when Internet-based data collection was initially based on RDD recruiting (as in the KN sample). In comparing unweighted data against U.S. Census Bureau demographic characteristics, there were no clear differences between the telephone and KN samples. Finally, respondents' lack of interest or familiarity with a topic can be a source of relatively low response rates. However, some recent research (Pineau and Slotwiner, 2004) indicates that increasing the response rate does not have a significant effect on analysis outcomes if the nonresponse rate is less than 80%, as in our case.

Table 1. Descriptive Statistics for Budget Constraint Questions

	Means, by Survey Mode [N = overall survey responses, and n = usable responses to question]	
Budget Question	Telephone (TEL) Sample [N = 865]	Knowledge Networks (KN) Sample [N = 1,050]
QA: Percentage of household income available for total discretionary uses per month	18.62% [n = 791]	17.29% [n = 960]
QB: Percentage of household income available for environmental causes per month	1.66% [n = 705]	2.07% [n = 947]

Notes: The nonresponse rate $((1 - (n/N)) * 100)$ for QA and QB for the TEL sample is 8.5% and 18.5%, respectively; and 8.6% and 9.2%, respectively, for the KN sample. One possible reason for no comparable drop in response rate on QB for the KN sample is because QA and QB were presented on one webpage. While we use the notion of mental accounting to motivate this exploration, we do not know the exact mental account(s) individual respondents may be using (i.e., whether and what they consider discretionary income or income available for environmental causes). So their answers, and choice to respond, to questions QA and QB may depend on their status, interests, and familiarity with the budgeting process, as well as their understanding of the questions.

The payment amounts (PAY) were randomly allocated within a set of t ($) = {6, 12, 25, 75, 150, 225, 300, 500, 700, 900, 1,200, 1,800, and 2,400} dollars.

Modeling Considerations and Hypotheses

Recoding Voting Responses for Uncertainty

Respondents' certainty levels in the referendum voting decision may differ greatly because of differences in the strength of their preferences, familiarity with the Kyoto Protocol, etc. Further, Berrens et al. (2002) and others have identified an apparent pattern in dichotomous choice CV formats where "No" means "No," and "Yes" often means only "Maybe." Given these concerns and potential yea-saying (Blamey, Bennett, and Morrison, 1999) or upward hypothetical bias, we follow the general approach of Champ and Bishop (1997, 2001), Loomis and Ekstrand (1998), Berrens et al. (2002), and others in recoding selected Yes votes. Such recoding is based on responses to a follow-up question concerning the certainty level of their answers to the referendum question. That is, after respondents answered the advisory referendum question, a follow-up question was asked as to how certain their voting decisions were on a scale from 0 to 100, with 0 meaning absolutely certain of voting against it and 100 meaning absolutely certain of voting for it. Two separate recoding strategies are implemented in the study. A first recoding threshold is set at 50 (R50); i.e., only the Yes votes with a score of 50 and above were treated as Yes votes (and all other votes are No). This mild recoding is expected to filter any "noise" caused by confusion or misunderstanding (e.g., a Yes vote that is more certain of being a No gets recoded as a No). A second (separate) recoding threshold is set at 70 (R70), so that only relatively certain Yes votes are included; i.e., only Yes votes with a score of 70 and above were treated as Yes votes (all other votes are No).

Explanatory Variables

Explanatory variables used in various models are classified into two groups: socio-economic variables, and Kyoto Protocol-related variables. The socioeconomic variables include respondent's annual household income (*LINC*), education level (*EDUC*), ideology (*IDEO*), respondent' age (*AGE*), and indicator variables of respondent's gender (*MALE*) and environmental group membership (*MEMBER*). Variables relevant to the Kyoto Protocol include two attitudinal variables, respondent's attitude toward international treaties (*INTRTY*) and environmental problems (*BRINK*); one knowledge variable, respondent's knowledge about greenhouse gases (*GRGAS*); and two assessment variables, the effectiveness (*CONFID*) and the fairness (*FAIR*) of the Kyoto Protocol. Detailed definitions and the associated sample statistics are provided in table 2, partitioned by samples (TEL and KN) and treatments (*MA* = 0 and *MA* = 1).

We expect a priori that explanatory variables of *LINC, INTRTY, BRINK*, and *GRGAS* will positively influence the voting decisions—because the Kyoto Protocol is an international treaty (related to *INTRTY*) of limiting greenhouse gas emissions (related to *GRGAS*) to handle an important global environmental problem (related to *BRINK*). The more the respondents understand the Kyoto Protocol, and the greater their concern about environmental problems, the more likely they would vote for the treaty. We also anticipate the more positively the respondents assess the Kyoto Protocol (*CONFID* and *FAIR*), the more likely they would favor the treaty.

WTP Modeling

The modeling approach follows the conventional referendum CV model of Cameron and James (1987) to directly estimate a household WTP function (separately for both raw and recoded data). We begin with the underlying WTP function:

$$(1) \qquad\qquad WTP_i = f(\mathbf{x}_i, \beta, \sigma, \varepsilon_i) = e^{\beta'\mathbf{x}_i + \sigma\varepsilon_i},$$

where $\mathbf{x}_i$ is a vector of the selected explanatory variables of respondent i, β is the estimated coefficient of corresponding explanatory variables, σ is a variance parameter, and ε_i is a random error component with mean zero.

Since we cannot detect WTP responses directly in the referendum format, the latent function of an individual's true WTP can be observed by a discrete indicator variable, W_i, where

$$(2) \qquad\qquad W_i = 1 \text{ if } WTP_i \geq \mathbf{t}_i; \quad W_i = 0 \text{ otherwise,}$$

and $\mathbf{t}_i$ is the payment amount randomly assigned to respondent i. Thus, the probability of a Yes response is specified as:

$$(3) \qquad \Pr(Yes) = \Pr(W_i = 1) = \Pr(WTP_i > \mathbf{t}_i) = 1 - \Phi((\mathbf{t}_i - \beta'\mathbf{x}_i)/\sigma).$$

Based on the assumption of a lognormal distribution of the error term ε_i, a probit model is employed, and the log-likelihood function is:

$$(4) \qquad\qquad \log L = \Sigma\Big\{W_i\log\big[1 - \Phi((\log(\mathbf{t}_i) - \beta'\mathbf{x}_i)/\sigma)\big]$$
$$+ (1 - W_i)\log\big[\Phi((\log(\mathbf{t}_i) - \beta'\mathbf{x}_i)/\sigma)\big]\Big\}.$$

Table 2. Descriptive Statistics: Means and (Standard Deviations) for Selected Variables, by MA Treatment

| Variable | Description | TEL Sample [N = 796] | | KN Sample [N = 1,225] | |
		MA = 0 [N = 410]	MA = 1 [N = 386]	MA = 0 [N = 635]	MA = 1 [N = 590]
MA	Indicator variable of mental accounts treatment: 1 = mental accounts; 0 = only received standard reminder treatment	0.00 (0.00)	1.00 (0.00)	0.00 (0.00)	1.00 (0.00)
PAY	Randomly assigned payment amount ($), ranging from: 6, 12, 25, 75, 150, 225, 300, 500, 700, 900, 1,200, 1,800, and 2,400 dollars	610.64 (739.97)	702.39 (745.84)	593.63 (683.83)	645.45 (750.75)
EDUC	Education level indicator variable, 1–7 scale: 1 = less than high school, 5 = college graduate, 7 = completed some or all graduate school	4.23 (1.34)	4.38 (1.33)	3.84 (1.26)	3.99 (1.33)
AGE	Respondent's age in years/100	0.42 (0.15)	0.42 (0.15)	0.47 (0.16)	0.45 (0.16)
MALE	Respondent's gender: 1 = male; 0 = female	0.54 (0.50)	0.49 (0.50)	0.54 (0.50)	0.52 (0.50)
BRINK	Respondent's attitude toward the relationship between environmental threats and human civilization, 0–10 scale: 0 = no real threat to civilization; 10 = human civilization is on the brink of collapse due to environmental threats	5.71 (2.07)	5.94 (1.97)	5.43 (2.24)	5.46 (2.19)
INTRTY	Respondent's view of international treaties as a way to handle environmental problems, 0–10 scale: 0 = very bad idea; 10 = very good idea	7.47 (2.72)	7.17 (2.76)	7.23 (2.62)	7.15 (2.61)
CONFID	Respondent's assessment of the effectiveness of the Kyoto Protocol, 0–1 scale: 0 = certain of no effect; 10 = certain to reduce global warming	6.18 (2.64)	5.95 (2.49)	5.75 (2.32)	5.58 (2.28)
GRGAS	Whether the respondent believes greenhouse gases cause average global temperatures to rise: 0 = no; 1 = yes	0.84 (0.36)	0.83 (0.38)	0.84 (0.37)	0.84 (0.37)
FAIR	Perceived fairness of the Kyoto Protocol, 0–10 scale: 0 = completely unfair; 10 = completely fair	4.51 (2.91)	4.64 (2.95)	4.92 (2.70)	4.78 (2.76)
IDEO	Political ideology index, 1–7 scale: 1 = strongly liberal; 7 = strongly conservative	4.28 (1.57)	4.30 (1.57)	4.09 (1.54)	4.03 (1.53)
MEMBER	Indicator variable for whether respondent is member of an environmental group: 1 = yes; 0 = no	0.11 (0.31)	0.14 (0.34)	0.06 (0.25)	0.07 (0.26)
LINC	Logarithm term of respondent's annual household income ($1,000s)	3.77 (0.80)	3.82 (0.79)	3.69 (0.71)	3.73 (0.71)
VYES	Indicator variable for whether respondent would vote for Senate ratification of Kyoto Protocol: 1 = yes; 0 = no	0.67 (0.47)	0.57 (0.50)	0.57 (0.50)	0.51 (0.50)
CERT	Respondent's reported certainty level of his/her voting decision, 0–100 scale: 0 = absolutely certain to vote against the Kyoto Protocol; 100 = absolutely certain to vote for the Kyoto Protocol	56.46 (38.04)	50.38 (37.82)	56.12 (35.11)	50.45 (35.70)

For the estimated models, goodness-of-fit is measured by McFadden's likelihood-ratio index (LRI) (McFadden, 1974; Greene, 2000).

For the models using recoded voting responses, we construct another index variable W_i'. The respondent must answer Yes to the referendum question ($W_i = 1$) and provide a follow-up certainty level (C_i) which is greater than the threshold certainty value of 50 or 70 in order for $W_i' = 1$. Then, parallel to the DC model using the raw voting responses, the individual's WTP must be inferred through the recoded indicator W_i':

$$(5) \qquad W_i' = \begin{cases} 1 \text{ if } WTP \geq 1 \text{ and } C_i \geq R; \\ 0 \text{ otherwise,} \end{cases}$$

where R is the recoding threshold (R50 or R70). Given the same distribution of the error term, the log-likelihood function of the recoded model is the same as in (4).

Finally, because mean WTP estimates can be very sensitive to outliers and the distributional assumption, we focus on the more robust median WTP, where median $WTP_i = \exp(\beta'\bar{x}_i)$, and x_i is replaced with its mean, $\bar{x}$ (Hanemann and Kanninen, 1999). The standard error of median WTP is calculated using the delta method (Greene, 2000).

Hypotheses

A series of hypotheses can now be put forward. First, prior to examining the WTP models, we investigate whether the standard reminder ($MA = 0$) versus the mental account treatment ($MA = 1$) (i.e., absence or presence of the two questions on discretionary budget) is significantly related to a respondent's voting decision. Hypothesis H_1 is tested against the null of no effect:

■ H_1: $MA = 0$ vs. $MA = 1$. *The mental account treatment is significantly related to the voting decision.*

A nonparametric Mantel-Hanzael χ^2 test is implemented, and it is expected that the evidence will support H_1 (a significant treatment effect).

Based on our WTP modeling strategies, the following hypotheses are tested across the models using both raw and recoded responses (R50 and R70, separately). To start, we directly test for a mental account treatment effect (the presence of a latent structural break between vectors of estimated coefficients under the $MA = 0$ control group and the $MA = 1$ treatment group). Hypothesis H_2 is tested against the null that the vectors of coefficients are the same, i.e., $\beta_i^{MA=0} = \beta_i^{MA=1}$:

■ H_2: $\beta_i^{MA=0} \neq \beta_i^{MA=1}$. *The coefficients on MA = 0 and MA = 1 are not equal.*

To examine whether a mental accounts treatment effect exists, a likelihood-ratio test is applied. We expect to reject the null.

Assuming a significant mental accounts treatment effect is found (i.e., the evidence supports H_1 and H_2), then the question is whether this translates into *significant* differences in median annual household WTP estimates. Let $WTP^{MA=0}$ and $WTP^{MA=1}$ be the median WTP under the $MA = 0$ and $MA = 1$ treatments, respectively. Then the final hypothesis (against the null hypothesis of no difference) tested is:

■ H_3: $WTP^{MA=1} < WTP^{MA=0}$.

The expectation is that the treatment causes respondents to more directly consider discretionary income and any mental accounts, potentially reducing any yea-saying or upward hypothetical bias, and that the estimated median $WTP^{MA=1}$ will be significantly smaller than $WTP^{MA=0}$.

Results

Measurement of Association Regarding Mental Accounts Effect

Before proceeding to the WTP modeling results, we investigate whether there exists an overall difference in the proportions of No votes between the standard reminder and the mental accounts treatments. The payment amount was randomly assigned and is negatively (positively) related to the proportion of Yes (No) votes for TEL, KN, and pooled samples across the raw and recoded data (R50 and R70). Further, along with using the median WTP estimator, using the recoded responses helps to mitigate a "fat tail" problem (high acceptance rate at high payment amounts). For example, the percentage of Yes responses at $2,400 was 35% for the raw data, it drops to 29% for R50 responses, and further down to 25% for R70 responses.

To examine whether the treatment affected voting responses across various payment amounts, the TEL and KN data sets are post-stratified into the standard reminder only ($MA = 0$) and mental accounts ($MA = 1$) samples. Then we measure the relationship between the standard reminder versus mental accounts treatment and respondent's voting behavior while controlling for the level of the payment amount. The Mantel-Hanzael χ^2 values are 6.46 and 4.34 for the TEL and KN data sets, respectively, which are both significant at the 0.05 level. Therefore, there is a statistically significant association between the presence of the mental accounts treatment and voting responses for these individual data sets. Further, confidence intervals of the odds ratios are significantly greater than 1, which indicates the respondents who received only standard reminders ($MA = 0$) are more likely to vote for the treaty. Not surprisingly, these relationships are more explicit for the pooled sample. The Mantel-Hanzael χ^2 value is 9.79, and is significant at the 0.01 level. Thus, results support H_1. On average, respondents facing the mental accounts treatment ($MA = 1$) are less likely to vote Yes, while controlling for the level of the payment amount, and across both survey modes.

Structural Break Tests Between Treatment Samples ($MA = 0$ versus $MA = 1$)

As developed previously, we estimate log-probit models using the raw data and two recodings (R50 and R70). For each model, a dummy variable (MA) is incorporated to indicate the treatment. $MA = 0$ and $MA = 1$ are then estimated separately to further compare WTP measurements. The results of these estimations are presented in table 3.

Before turning to the structural break tests, one can also observe a general consistency in the results for the explanatory variables across the various models in table 3. For example, the estimated coefficient on the respondent's confidence that the Kyoto Protocol would be effective ($CONFID$) is positive and significant in all three models, suggesting that the more respondents believe the treaty would reduce global warming,

Table 3. Estimation Results of Lognormal Models (with asymptotic standard errors in parentheses)

Variable	MODEL		
	Raw Data [N = 2,021]	R50 Recoded Data [N = 2,021]	R70 Recoded Data [N = 2,021]
Intercept	-11.35***	-9.93***	-11.48***
	(1.87)	(1.72)	(1.91)
EDUC	0.41***	0.44***	0.38***
	(0.13)	(0.12)	(0.13)
AGE	0.16	-0.10	0.72
	(0.99)	(0.94)	(1.00)
MALE	0.73**	0.69**	0.77**
	(0.31)	(0.30)	(0.32)
BRINK	0.30***	0.27***	0.28***
	(0.09)	(0.08)	(0.09)
INTRTY	0.50***	0.40***	0.51***
	(0.08)	(0.08)	(0.08)
CONFID	0.88***	0.82***	0.89***
	(0.11)	(0.10)	(0.11)
GRGAS	2.24***	2.08***	2.24***
	(0.53)	(0.49)	(0.54)
FAIR	0.24***	0.26***	0.23***
	(0.06)	(0.06)	(0.06)
IDEO	-0.46***	-0.37***	-0.50***
	(0.11)	(0.10)	(0.11)
MEMBER	1.34**	1.49***	1.46***
	(0.54)	(0.52)	(0.55)
LINC	0.66***	0.65***	0.66***
	(0.22)	(0.21)	(0.23)
KN	-0.64*	-0.59*	-0.76**
	(0.32)	(0.31)	(0.33)
MA	**-1.09***	**-0.97***	**-0.99***
	(0.32)	**(0.30)**	**(0.32)**
σ_{WTP}	4.67***	4.51***	4.73***
	(0.40)	(0.37)	(0.41)
— Summary Statistics —			
-2 LnL	2,521.1	2,034.5	1,987.5
-2 LnL^{Restricted} ($\beta_{Slopes} = 0$)	3,192.3	2,790.7	2,755.5
McFadden's LRI	0.21	0.29	0.28
Median WTP ($) (standard error)	474.14*** (59.89)	92.88*** (14.98)	44.06*** (8.91)

(continued . . .)

Table 3. Continued

	MODEL		
Variable	Raw Data [N = 2,021]	R50 Recoded Data [N = 2,021]	R70 Recoded Data [N = 2,021]
	— Comparison Statistics for MA = 1 and MA = 0 —		
Median $WTP^{MA=1}$ ($)	327.15***	59.36***	28.56***
(standard error)	(58.28)	(15.24)	(9.38)
Median $WTP^{MA=0}$ ($)	647.50***	148.58***	71.18***
(standard error)	(116.90)	(30.97)	(17.67)
t-Value (Median $WTP^{MA=1}$ vs. $WTP^{MA=0}$)	−2.45***	−2.58***	−2.13**
χ^2 structure ($\beta^{MA=1}$ vs. $\beta^{MA=0}$)	34.00***	30.70***	35.00***

Notes:

- Single, double, and triple asterisks (*) denote statistical significance at the 10%, 5%, and 1% levels, respectively.
- Average $\ln\hat{L}_0$ is the logarithm likelihood value of restricted function where all slope parameters equal zero and free constant only.
- McFadden's likelihood ratio index (LRI) = $1 - \ln\hat{L}_u/n\hat{L}_0$.
- Median $WTP^{MA=1}$ and $WTP^{MA=0}$ are calculated using samples $MA = 1$ and $MA = 0$, respectively, for raw data, R50, and R70 recoded samples.
- The value of the χ^2 structure test is calculated using $MA = 1$ and $MA = 0$ samples, respectively, for raw data, R50, and R70 recoded data.

the more they would pay for ratification of the Kyoto Protocol. The estimated coefficient on the attitudinal variable *GRGAS* is also positive and significant in all three models. That is, the more the respondents believe the greenhouse gases cause global tempera-ture to rise, the higher their WTP for supporting the Protocol. Further, higher household income (*LINC*), and being an environmental group member (*MEMBER*) will lead to a higher probability of voting Yes, and WTP for supporting the Protocol.

Different sets of estimated coefficients are used to investigate the mental account treatment effect (structural break tests). As reported in the bottom section of table 3, the results of the log-likelihood ratio test are 34, 31, and 35 for the raw, R50, and R70 samples, respectively, and χ^2 statistics are all significant at the 0.01 level. Results indi-cate an overall structural break in the $MA = 0$ and $MA = 1$ samples. The MA treatment influences respondents' voting decision substantially. The results are also supported by the estimated coefficient of MA (β_{MA}), which is significant at the 0.01 level in all three models.[6] Thus, evidence from the WTP models shows that respondents in the treatment group ($MA = 1$) have significantly different voting behavior.

Comparison of Estimated Median WTP

We now want to examine whether the significant treatment effect translates into signif-icant differences in median WTP estimates (bottom section of table 3). For $MA = 0$, respondents are simply reminded of their budget in general. Results indicate that esti-mated $WTP^{MA=0}$ values are always larger relative to values from models using different

---

[6] We also estimated a set of matched models using TEL and KN samples separately. The quantitative results are largely consistent with the pooled sample estimations presented here. Results are available from the lead author upon request.

budget constraints for all of the three models. Since β_{MA} is significantly negative, the mental account treatment is expected to produce smaller WTP estimates—but the interesting issue is the magnitude and significance of such differences. For the model using the raw data, the median annual household WTP for the U.S. Senate ratification of the Kyoto Protocol for the treatment group ($327) is 50% smaller than for the control group ($647). For the models using the recoded responses, the median annual household WTP for the treatment groups is even lower (60%) for the ratification by the control group ($59 vs. $148 for R50, and $29 vs. $71 for R70). The one-tailed t-test results support H_3 (0.05 level). Across the three models (for the raw, R50, and R70 samples), the average median $WTP^{MA=1}$ ranges from approximately 50% to 60% less than median $WTP^{MA=0}$. In summary, results suggest a mental account treatment—asking respondents about their discretionary income—would decrease estimated median WTP significantly.[7]

Discussion and Conclusions

As with assessing other significant changes for complex environmental goods, consideration of the economic costs and benefits of reducing global climate change will continue to be an important informational input to planning and decision making. To the extent that bringing the United States back into serious international negotiations may greatly impact the efficacy of the Kyoto Protocol (or some variant thereof), then information on U.S. household preferences remains policy relevant.

This study uses a unique combination of mixed-mode (telephone and Internet) samples to explore the effects of a mental accounts-type treatment on U.S. household referendum voting (and annual WTP) for ratification of the Kyoto Protocol. The treatment consists of two questions related to discretionary income (respondents are also provided with a standard budget reminder immediately prior to the referendum questions). In the control group, respondents are only provided with the standard budget reminder of their household income. Results of our study suggest that the presence of a mental accounts treatment, where respondents are asked to provide two categories of discretionary income, has a significant effect on voting responses. Nonparametric tests suggest the treatment significantly lowers the probability of Yes votes on the advisory referendum. Further, the WTP modeling results indicate the presence of a structural break under the treatment; using the mental account information ($MA = 1$ model) always produces a significantly lower estimate of median annual household WTP (ranging from 50% to 60%, across the raw and recoded data).

From a methodological perspective, using mental accounting-type questions appears to hold considerable promise for mitigating potential yea-saying or upward hypothetical bias. As a final caveat, in an extended version of this paper (available from the authors upon request, and see Li, 2003), we have explored using the monetary information on budget proportions from the treatment questions (QA and QB) explicitly in WTP modeling. These efforts include implementing the Beta model (Haab and McConnell, 1998) where proportional budget information can be directly used, and ad hoc restrictions on the upper limit of integration for the WTP distribution. However, such efforts

---

[7] We also compared the estimated median $WTP^{MA=0}$ and $WTP^{MA=1}$ based on individual sample means of $MA = 0$ and $MA = 1$, respectively. The t-statistics again support H_3 (0.01 level, one-tailed t-test). Results are available from the lead author upon request.

are limited in that we do not know the exact mental account (e.g., total income, discretionary income, income available for environmental causes, etc.) any individual may be using in the referendum vote (with the given payment vehicle). Thus, we focus here on strict tests of the effect of introducing the treatment. Future research may be extended by elicitations that explicitly map both the respondent-articulated mental account and a valuation response.

[*Received February 2005; final revision received May 2005.*]

References

Arrow, K., R. Solow, P. Portney, E. Leamer, R. Radner, and H. Schuman. "Report of the NOAA Panel on Contingent Valuation." *Federal Register* 15(1993):4602–4614.

Baron, J. "Confusion of Relative and Absolute Risk in Valuation." *J. Risk and Uncertainty* 14,2(1997): 301–309.

Bateman, I., and I. Langford. "Budget Constraint, Temporal, and Question-Ordering Effects in Contingent Valuation Studies." *Environment and Planning* A29(1997):1215–1228.

Bateman, I., and K. Willis, eds. *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the U.S., EU, and Developing Countries.* London: Oxford University Press, 1999.

Berrens, R., A. Bohara, H. Jenkins-Smith, C. Silva, and D. Weimer. "The Advent of Internet Surveys for Political Research: A Comparison of Telephone and Internet Samples." *Political Analysis* 11,1(2003):1–23.

———. "Information and Effort in Contingent Valuation Surveys: Application to Global Climate Change Using National Internet Samples." *J. Environ. Econ. and Mgmt.* 47,2(2004):331–363.

Berrens, R., H. Jenkins-Smith, A. Bohara, and C. Silva. "Further Investigations of Voluntary Contribution Contingent Valuation: Fair Share, Time of Contribution, and Respondent Uncertainty." *J. Environ. Econ. and Mgmt.* 44,1(2002):144–168.

Bishop, R. "Where to from Here?" In *A Primer on Nonmarket Valuation*, eds., P. Champ, K. Boyle, and T. Brown, pp. 537–567. London: Kluwer Academic Publishers, 2003.

Blamey, R. K., J. W. Bennett, and M. D. Morrison. "Yea Saying in Contingent Valuation Surveys." *Land Econ.* 75,1(1999):126–141.

Cameron, T. A., and M. D. James. "Efficient Estimation Methods for 'Closed Ended' Contingent Valuation Surveys." *Rev. Econ. and Statis.* 69,2(1987):269–276.

Champ, P. A., and R. C. Bishop. "Using Donation Mechanisms to Value Nonuse Benefits from Public Goods." *J. Environ. Econ. and Mgmt.* 33(1997):151–162.

———. "Donation Payment Mechanisms and Contingent Valuation: An Empirical Study of Hypothesis Bias." *Environ. and Resour. Econ.* 19(2001):383–402.

Deaton, A., and J. Muellbauer. *Economics and Consumer Behavior.* London: Cambridge University Press, 1980.

Duborg, W., M. Jones-Lee, and G. Loomes. "Imprecise Preferences and Survey Design in Contingent Valuation." *Economica* 64(1997):681–702.

Greene, W. H. *Econometric Analysis*, 4th ed. Englewood Cliffs, NJ: Prentice-Hall, 2000.

Haab, T., and K. McConnell. "Referendum Models and Economic Values: Theoretical, Intuitive, and Practical Bounds on Willingness to Pay." *Land Econ.* 74,2(1998):216–229.

Hailu, A., W. Adamowicz, and P. C. Boxall. "Complements, Substitutes, Budget Constraints, and Valuation." *Environ. and Resour. Econ.* 16(2000):51–68.

Hanemann, W. H., and B. Kanninen. "The Statistical Analysis of Discrete-Response CV Data." In *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the U.S., EU, and Developing Countries*, eds., I. Bateman and K. Willis, pp. 302–441. London: Oxford University Press, 1999.

Hoevenagel, R. "An Assessment of Contingent Valuation Surveys." In *Pricing the European Environment*, ed., S. Navrud, pp. 177–194. Oslo/Oxford: Scandinavian University Press, 1992.

Kemp, M., and C. Maxwell. "Exploring a Budget Context for Contingent Valuation Estimates." In *Contingent Valuation: A Critical Assessment*, ed., J. A. Hausman, pp. 217–265. Amsterdam: North-Holland Press, 1993.

Kotchen, M., and S. Reiling. "Do Reminders of Substitutes and Budget Constraints Influence Contingent Valuation Estimates? Another Comment." *Land Econ.* 75,4(1999):478–482.

Krosnick, J., and L. Chang. "A Comparison of the Random Digit Dialing Telephone Survey Methodology with Internet Survey Methodology as Implemented by Knowledge Networks and Harris Interactive." Unpub. manu., Dept. of Psychology and Political Science, Ohio State University, 2001.

Landler, M. "Mixed Feelings as Treaty on Greenhouse Gases Takes Effect." *New York Times* (16 February 2005). Online. Available at www.nytimes.com/2005/02/16/business/worldbusiness/16kyoto.html?pagewanted=1&ei=5070&en=f4f4c54b8cda9573&ex=1109221200.

Li, H. "Explorations in Valuing Global Climate Change: Testing Treatment Effects with Internet and Telephone Samples." Unpub. Ph.D. diss., Dept. of Econ., University of New Mexico, Albuquerque, December 2003.

Li, H., R. Berrens, A. Bohara, H. Jenkins-Smith, C. Silva, and D. Weimer. "Would Developing Country Commitments Affect U.S. Households' Support for a Modified Kyoto Protocol?" *Ecological Econ.* 48(2004):329–343.

List, J., R. Berrens, A. Bohara, and J. Kerkvliet. "Examining the Role of Social Isolation on Stated Preferences." *Amer. Econ. Rev.* 94,3(2004):741–752.

Little, J., and R. Berrens. "Explaining Disparities Between Actual and Hypothetical Stated Values: Further Investigation Using Meta-Analysis." *Econ. Bulletin* 3,6(2004):1–13.

Loomis, J., and E. Ekstrand. "Alternative Approaches for Incorporating Respondent Uncertainty When Estimating Willingness to Pay: The Case of the Mexican Spotted Owl." *Ecological Econ.* 27(1998): 29–41.

Loomis, J., A. Gonzalez-Caban, and R. Gregory. "Do Reminders of Substitutes and Budget Constraints Influence Contingent Valuation Estimates?" *Land Econ.* 70(1994):499–506.

———. "Do Reminders of Substitutes and Budget Constraints Influence Contingent Valuation Estimates? Reply." *Land Econ.* 71,4(1995):544–545.

Magnussen, K. "Valuation of Reduced Water Pollution Using the Contingent Valuation Method— Testing for Mental Accounts and Amenity Misspecification." In *Pricing the European Environment*, ed., S. Navrud, 194–230. Oslo/Oxford: Scandinavian University Press, 1992.

McFadden, D. "The Measurement of Urban Travel Demand." *J. Public Econ.* 3(1974):303–328.

Mitchell, R., and R. T. Carson. "Using Surveys to Value Public Goods: The Contingent Valuation Method." Washington, DC: Resources for the Future, 1989.

Moon, P., K. Keasey, and D. Duxbury. "Mental Accounting and Decision Making: The Relationship Between Relative and Absolute Savings." *J. Econ. Behavior and Organization* 38,2(1999): 145–153.

Pineau, V., and D. Slotwiner. "Probability Samples v. Volunteer Respondents in Internet Research: Defining Potential Effects on Data and Decision-Making in Market Applications." Research report, Knowledge Networks, Inc., 2004. Online. Available at http://www.knowledgenetworks.com/info/press/papers/ Volunteer%20white%20paper%2011-19-03.pdf.

Smith, V. K. "Comments on 'Exploring a Budget Context for Contingent Valuation Estimates.'" In *Contingent Valuation: A Critical Assessment*, ed., J. A. Hausman, p. 266. Amsterdam: North-Holland Press, 1993.

Thaler, R. "Mental Accounting and Consumer Choice." *Marketing Sci.* 4(1985):199–214.

———. "Anomalies: Savings, Fungibility, and Mental Accounts." *J. Econ. Perspectives* 4(1990):193–205.

———. "Mental Accounting Matters." *J. Behavioral Decision Making* 12,3(1999):183–206.

United Nations. Online press release. Kyoto Protocol taking effect on its 128 parties. UN website, February 16, 2005. Online at http://unfccc.int/files/press/news_room/press_releases_and_advisories/application/pdf/press041118_eng.pdf.

Whitehead, J., and G. Blomquist. "Do Reminders of Substitutes and Budget Constraints Influence Contingent Valuation Estimates? Comment." *Land Econ.* 71(1995):541–543.

————. "Do Reminders of Substitutes and Budget Constraints Influence Contingent Valuation Estimates? Reply to Another Comment." *Land Econ.* 75(1999):483–484.

Willis, K., and G. Garrod. "Valuing Landscape: A Contingent Valuation Approach." *J. Environ. Mgmt.* 37(1993):1–12.