Detection and Mitigation of Hypothetical Bias in Contingent Valuation With An Application To Curbside Recycling

David Aadland
Utah State University

Arthur J. Caplan
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/eri

Recommended Citation
https://digitalcommons.usu.edu/eri/222
DETECTION AND MITIGATION OF HYPOTHETICAL BIAS IN CONTINGENT VALUATION WITH AN APPLICATION TO CURBSIDE RECYCLING

by

DAVID AADLAND

ARTHUR J. CAPLAN

Department of Economics
Utah State University
3530 Old Main Hill
Logan, UT 84322-3530

February 2001
DETECTION AND MITIGATION OF HYPOTHETICAL BIAS
IN CONTINGENT VALUATION WITH AN APPLICATION
TO CURBSIDE RECYCLING

David Aadland, Assistant Professor
Arthur J. Caplan, Assistant Professor

Department of Economics
Utah State University
3530 Old Main Hill
Logan, UT 84322-3530

The analyses and views reported in this paper are those of the author(s). They are not necessarily endorsed by the Department of Economics or by Utah State University.

Utah State University is committed to the policy that all persons shall have equal access to its programs and employment without regard to race, color, creed, religion, national origin, sex, age, marital status, disability, public assistance status, veteran status, or sexual orientation.

Information on other titles in this series may be obtained from: Department of Economics, Utah State University, 3530 Old Main Hill, Logan, Utah 84322-3530.

Copyright © 2001 by David Aadland and Arthur J. Caplan. All rights reserved. Readers may make verbatim copies of this document for noncommercial purposes by any means, provided that this copyright notice appears on all such copies.
DETECTION AND MITIGATION OF HYPOTHETICAL BIAS IN CONTINGENT VALUATION WITH AN APPLICATION TO CURBSIDE RECYCLING

David Aadland and Arthur J. Caplan

ABSTRACT

In this paper, we use a unique curbside-recycling data set to test the effectiveness of “cheap talk” and “preference uncertainty” in mitigating hypothetical bias in contingent valuation. The sample includes two types of households—those located in communities with curbside recycling programs (mandatory or voluntary) and those in communities without curbside recycling. Using stated and revealed preference data, detect significant hypothetical bias. Cheap talk and preference-uncertainty controls are partially effective in mitigating the bias.

JEL Classification: C35, D12
DETECTION AND MITIGATION OF HYPOTHETICAL BIAS
IN CONTINGENT VALUATION WITH AN APPLICATION
TO CURBSIDE RECYCLING

INTRODUCTION

There is an active debate regarding the reliability of information from contingent valuation (CV), and more generally, regarding the issue of hypothetical bias from stated-preference methods (Diamond and Hausman, 1994; Carson et al., 2000; Arrow et al., 1993). Although there is disagreement about the extent of the bias, nearly all researchers agree that it exists. Hypothetical bias arises when people are asked to hypothetically state or select a maximum amount they are willing to pay for a good or service, even though they know they will never have to actually pay for it. Researchers have typically used two general methods to detect and mitigate this bias. The first involves innovations in the design of the survey instrument. The second involves combining results of actual market outcomes (i.e., revealed preference (RP) data) with stated preference (SP) data to improve the reliability of the resulting welfare estimates.

Recent innovations in survey design to mitigate hypothetical bias include: (1) reminder statements of substitutes, budget constraints, or the hypothetical nature of the good in question (Neill et al., 1994; Neill, 1995; Loomis et al., 1996; Cummings and Taylor, 1999; List, 2001), (2) use of referendum formats (Mitchell and Carson, 1989; Loomis, 1990; Arrow et al., 1993; Alberini, 1995b; Boyle et al., 1996; Bohara et al., 1998; Welsch and Poe, 1998; Carson et al., 2000), and (3) follow-up questioning (Cameron and Quiggin, 1994; Alberini, 1995b; Kanninen, 1995; Li and Mattsson, 1995; Blumenschein et al., 1998; Berrens et al., 2000; Champ et al., 1997). In this paper, we test two of these innovations.

The first is the use of “cheap talk” in CV, as introduced by Cummings and Taylor (1999).1 Cheap talk is information provided prior to the willingness-to-pay (WTP) questions reminding respondents that they are being asked to value a hypothetical program. Essentially, it is a scripted explanation of how the program’s hypothetical nature can bias WTP responses. Cummings and Taylor find some evidence that households receiving cheap talk will, on average, report lower WTP values than those that do not (i.e., that cheap talk may correct for upward bias in WTP responses), but they also suggest the need for further empirical testing. In particular,

---

1 Although Cummings and Taylor (1999) were the first to use the game-theoretic term “cheap talk” in the context of CV, Loomis et al. (1994) employed a similar approach in an earlier study. In a laboratory experiment, Loomis et al. were able to reduce the discrepancy between hypothetical and actual WTP by issuing a reminder to their subjects (prior to the WTP question) that “... although the question is hypothetical, we want you to answer as if it were real . . .”
they recommend reducing the length of the cheap-talk script to be more compatible with telephone applications of CV.²

The second innovation controls for preference uncertainty by asking respondents in a follow-up question how certain they are of their WTP responses. For example, in their study of forest protection in Sweden, Li and Mattsson (1995) develop a structural model that explicitly accounts for preference uncertainty. They find a significantly lower mean WTP, suggesting that ignoring preference uncertainty may upwardly bias estimates of WTP. Berrens et al. (2000) also use preference-uncertainty information in their study of in-stream flow protection in the Middle Rio Grande of New Mexico. However, rather than develop a new structural model, they attempt to reduce potential hypothetical bias through a re-coding of the WTP data.³ As with cheap talk, we incorporate preference-uncertainty measures directly into our survey design in an attempt to mitigate hypothetical bias associated with WTP for curbside recycling.

Regarding the second general method to mitigate hypothetical bias - using SP and RP data – the study closest in spirit to ours is Whitehead et al. (2000).⁴ Whitehead et al. (2000) find significant differences between stated and revealed preferences in their sample of current users and non-users of the Albemarle and Pamlico Sounds in North Carolina, indicating potential structural changes in trip demand (i.e., hypothetical bias in their SP data). They use intercept and slope dummy variables to control for the structural change associated with stated preferences for current and improved quality, respectively. By incorporating these controls, Whitehead et al. account for current non-participants who plan to participate in the future regardless of whether quality improvements are undertaken. The authors find that these variables are significant in explaining trip behavior in their sample, and therefore conclude that RP and SP data represent

² Cummings and Taylor (1999) mention two as yet unpublished studies that find no significant cheap talk effect using shortened scripts. List (2001) finds that the effectiveness of the long-script form of cheap talk depends upon respondent experience with the good being valued.
³ We think of hypothetical bias in broad terms, as any deviation of an individual’s stated WTP from their actual WTP that is due to the hypothetical nature of the transaction or the good itself. In this sense, post-decisional measures of uncertainty are identifying potential deviations in stated WTP from their true values. As a consequence, estimation methods, such as those in Li and Mattsson (1995) or Berrens et al. (2000), although they do not explicitly refer to the term “hypothetical bias,” are attempting to correct for a bias in stated WTP values that is related to the hypothetical nature of CVM framework.
⁴ There are a multitude of studies that use RP and SP data together to improve the reliability of the welfare estimates, e.g., Nestor, 1998; Loomis, 1993; Brookshire and Coursey, 1987; Bishop, Heberlein, and Kealy, 1983; and Bishop and Heberlein, 1979. Cameron (1992) and Adamowicz et al. (1994) combine CV responses with travel-cost information from the same individuals, who are all users of the environmental good. Cameron finds that combining SP and RP information enhances the overall reliability of welfare estimates. Adamowicz et al. obtain similar results. Both of these findings are consistent with Carson et al. (1996), whose meta-analysis shows that SP and RP welfare estimates are in general similar, therefore supporting the assumption of identical preference structures across SP and RP data.
the same underlying behavior at both current and improved quality levels after accounting for hypothetical bias.\(^5\)

Our data set consists of a sample of over 1000 households from the state of Utah who were asked to value either their existing curbside recycling program (CRP), or a hypothetical program if their community does not currently provide one. The nature of our data set is therefore similar to Whitehead et al.'s. It includes both RP and SP information from households that choose whether to participate in an actual CRP, as well as SP information from non-participant households that are either unaware that a CRP exists in their community or are residing in a community without a CRP.\(^6\) Because the attributes of the hypothetical CRP described to the non-participants are identical to those of an existing voluntary CRP (where households only pay if they use the service), we are able to estimate a sub-model using data solely from these two groups. This added feature of our data set enables us to detect hypothetical bias through a direct comparison of stated and revealed preferences. This is a direct RP-SP comparison because we have information on actual purchasing decisions of households rather than indirect information on preferences such as that provided by travel costs. We then purposefully include additional explanatory variables (i.e., cheap talk and preference uncertainty) into our survey design with the intent of mitigating the bias.

Similar to List (2001), we find evidence that cheap talk is effective in mitigating hypothetical bias for certain types of individuals, although the effect is not always precisely estimated. In particular, cheap talk appears to be effective for individuals who tend to state relatively higher WTP amounts and as a result may be more prone to hypothetical bias.

Regarding preference uncertainty, we create a post-decisional confidence measure through a follow-up debriefing question and then test the measure's effect on WTP in two ways – as an additional explanatory variable and through data re-coding. Similar to Champ et al. (1997), Blumenschein et al. (1998) and Berrens et al. (2000), we find that individuals who are uncertain about their WTP response state significantly higher WTP values than individuals who are very certain about their responses. By re-coding our data in such a way that respondents who are

---

\(^5\) In a similar paper using the same data set, Huang et al. (1997) investigate the conditions under which SP and RP recreation-demand data should be combined under the same assumed preference structure. Their results rely heavily on information about "ex-ante trips with current quality" from current resource users. They find that the consistency of joint SP/RP estimation of WTP for an improvement in quality requires use of ex-ante, rather than ex-post, trips with current quality.

\(^6\) Thus, in contrast with Whitehead et al., our data includes information from non-participants who are physically precluded from using the environmental good (e.g., if they are located in a community that does not currently have a CRP available).
uncertain about their preferences receive lower WTP, we are able to further reduce an apparent upward hypothetical bias in our CV estimates. 7

The next section presents the econometric model used to estimate our overall welfare measures, as well as the specific methods used to detect and mitigate hypothetical bias in our data. We then briefly discuss our survey instrument and the data, followed by a presentation of our empirical results. The final section concludes with a summary of our findings and policy recommendations.

METHODOLOGY AND ECONOMETRIC MODEL

This section is divided into three subsections. The first subsection discusses the general double-bounded dichotomous-choice (DBDC) model we use to obtain our welfare estimates. The second then introduces the explicit hypothesis we test to detect whether hypothetical bias exists in our SP data. The third subsection similarly discusses the hypotheses for cheap talk and preference uncertainty.

Econometric Model

Our econometric approach to estimating various WTP measures follows Cameron and James (1987). Their procedure has become commonplace in contingent-valuation studies and has been shown by McConnell (1990) to be dual to the utility-theoretic approach presented by Hanemann (1984), which specifies that the WTP equation is derived from a household's constrained utility-maximization problem. The Cameron and James approach has the advantage that various statistics such as the mean WTP estimates, marginal effects, standard errors, and test statistics are all easily accessible from the estimated WTP equation.

WTP questions set in the DBDC format elicit a household's WTP through a sequence of dichotomous-choice (i.e., yes or no) valuation questions. The first question is typically something like: "Would you be willing to pay $t for the service?" The opening bid $t is chosen randomly from a set of values that depends upon the good or service being valued. 8 By randomly choosing the opening bid, the possible effects of "starting-point bias" are reduced (Alberini, 1995a and 1995b; Arrow et al., 1993; and Cameron, 1988). Based on the response to

7 In addition to reducing potential upward bias, data re-coding enables us to estimate lower-bound welfare measures from our data (Champ et al., 1997).
8 In this study, opening bids were randomly chosen integers from $2 to $10, reflecting the approximate range of values for existing curbside recycling programs.
the opening bid, the respondent is then asked a follow-up question of the same nature, but with a larger bid, \( \tau_1 \), if she answered “yes” (i.e., willing to pay at least \( \tau \) for the service) or a smaller bid \( \tau_L \) if she answered “no” (i.e., unwilling to pay \( \tau \) for the service).

Based on the responses to both the opening bid and follow-up questions, we are able to place the respondent’s true (yet unknown) WTP in one of four regions: \((-\infty, \tau_L \), \((\tau_L, \tau) \), \((\tau, \tau_H) \) or \((\tau_H, \infty) \). For this project, we have set \( \tau_L = 0.5\tau \) and \( \tau_H = 2\tau \). Unlike other CV studies, we follow-up with the following third valuation question for those who respond “no” to the first two valuation questions: “Would you be willing to use the service if it were free of charge?” We ask this question based on previous experience with household recycling surveys, which suggests that some households apparently need to be paid (i.e., have negative WTP values) to participate (Aadland and Caplan, 1999; Haab and McConnell, 1997). As a result, our survey generates five rather than four valuation regions with \((-\infty, \tau_L \) being replaced by \((-\infty, 0) \) and \((0, \tau_L) \).

We next posit that the household's true WTP (WTP*) can be represented by the equation

\[
\text{WTP}_i^* = X_i\beta + \epsilon_i, \tag{1}
\]

where \( X_i \) is a row vector of household and community-specific control variables, \( \beta \) is a corresponding column vector of coefficients, and \( \epsilon_i \) is a normally distributed error term for households \( i = 1, \ldots, n \).\(^9\) We estimate two models that include households from communities with and without curbside service.

By assuming a normal distribution and independence across error terms, we then form the likelihood function conditional on (1) and the observed data. Letting \( \Phi \) indicate the mean-zero normal cumulative density function, we may write the probability that a given household \( i \)'s true WTP falls in each of the five intervals as:

\(^9\) Following McConnell (1990), we assume that WTP* follows from household \( i \)'s constrained utility maximization problem. If household \( i \) is either located in a community without a CRP, or located in a community with a CRP (voluntary or mandatory) which it does not use, then we can think of its minimum expenditure function \( m_i(q_0, X_i, u_i) \) as being the inverse of its indirect utility function \( u_i(q_0, X_i, I_i) \) with respect to income \( I_i \), where \( X_i \) is as defined in the text, \( q_0 = q_0 \) and \( q_1 = q_1 \) imply non-existence and existence, respectively, of a CRP, and the price vector for all other goods is suppressed. If the respondent answers yes to the question, "Would you willingly pay \$\tau_i \) per month for curbside recycling?" then \( \text{WTP}_i = I_i - m_i(q_i, X_i, u_i(q_0, X_i, I_i)) + \epsilon_i > \tau_i \), which for empirical purposes we express as (1), where vector \( X_i \) includes \( \tau_i \). If household \( i \) is located in a community with a CRP that it uses, then its minimum expenditure function is defined as \( m_j(p, X_i, u_j) \), where \( p \) is household \( i \)'s monthly fee. Now, if the respondent answers yes to the question, "Would you willingly pay \$\tau_i \) per month for curbside recycling?" then \( \text{WTP}_i = I_i - m_j(p, X_i, u_j(p_i, X_i, I_i)) + \epsilon_i > \tau_i \), where \( p_i \) is the personalized "choke" price at which household \( i \) would no longer willingly pay for the service. Again, for empirical purposes we express this result as equation (1).
\[ P_{1,i} = \text{Prob}(-\infty < WTP_i^* < 0) = \Phi(-X_i\beta); \]
\[ P_{2,i} = \text{Prob}(0 \leq WTP_i^* < 0.5\tau_i) = \Phi(0.5\tau_i - X_i\beta) - \Phi(-X_i\beta); \]
\[ P_{3,i} = \text{Prob}(0.5\tau_i \leq WTP_i^* < \tau_i) = \Phi(\tau_i - X_i\beta) - \Phi(0.5\tau_i - X_i\beta); \]
\[ P_{4,i} = \text{Prob}(\tau_i \leq WTP_i^* < 2\tau_i) = \Phi(2\tau_i - X_i\beta) - \Phi(\tau_i - X_i\beta); \]
\[ P_{5,i} = \text{Prob}(2\tau_i \leq WTP_i^* < \infty) = 1 - \Phi(2\tau_i - X_i\beta), \]

where \( \tau_i \) represents household \( i \)'s opening bid. Using (1) and (2), we can write the (log) likelihood function for all households in the sample as

\[ \ln(L) = \sum_{i=1}^{n} \sum_{j=1}^{5} \omega_{ij} \ln(P_{ij}), \]

where \( \omega_{ij} = 1 \) if the stated WTP value falls in the \( j^{th} \) region and 0 otherwise. Maximizing the (log) likelihood function (3) given normally distributed error terms, results in an estimation problem requiring nonlinear optimization techniques to generate estimates of the \( \beta \) parameters and the associated marginal effects (see Greene, 1999).\(^{10}\)

Recently, the "random effects" model as presented in Alberini et al. (1997) has become a common method for estimating models using referendum data with follow-up questions. The random-effects model allows for the possibility that individuals make errors in assessing their true WTP across the two valuation questions. As a result, an additional parameter measuring the correlation between latent WTPs associated with each question is introduced into the bivariate WTP distribution. Rather than estimate a random-effects model, we estimate the more traditional "double-bounded" WTP model for several reasons. First, we suspect that respondents valuing curbside recycling are less likely to be subject to these types of random WTP errors across questions. Curbside recycling is a good most households are familiar with and the described hypothetical CRP is relatively simple, placing few demands on the respondent such as recalling variations in attributes of the good being valued. Furthermore, Alberini et al. (1997)\(^{10}\)

\(^{10}\) Some of the respondents answered "Don't Know" to a one or more of the valuation questions. For these households, their unknown WTP does not fit into one of the five categories, but instead overlaps one or more of the intervals. For example, if a respondent answered "Don't Know" to whether they would be willing to pay $r and
find evidence that goods of this nature are less prone to these types of valuation errors across survey questions. Second, it is still an unsettled debate as to whether the bivariate approach to estimation gives more reliable welfare estimates than the more traditional univariate approach (Alberini, 1995a and Haab, 1998). Third, and finally, unlike other contingent valuation studies, we offer a second follow-up question to a subset of our respondents, thus generating the need for a trivariate WTP distribution, adding substantial complexity to the estimation procedure.

Detecting Hypothetical Bias

Typically CV researchers are unable to quantify the extent of hypothetical bias because they do not observe individuals revealing their preferences through actual purchasing decisions. However, in our case, we have observations from both individuals who have made a voluntary choice to either participate or not in a CRP, and from individuals who were asked to hypothetically place a value on a very similar CRP. The hypothetical program was described as follows:

... please imagine that you could have a service that regularly collects paper, plastic, glass, aluminum cans, tin cans, and cardboard. Your household would need to sort your recyclables (into groups of similar materials) and pay a fee for the recycling service, in addition to your current garbage collection fee.

To ensure that households were valuing similar CRPs in this comparison, we first isolated Salt Lake City (SLC) respondents who know that a CRP exists in their community and who stated that the fee for the existing program is approximately $2.00 per month. The CRP in Salt Lake City costs precisely $2.09, involves some sorting of recyclable materials, and picks up materials two to four times a month. Therefore, the SLC program is essentially identical to the hypothetical program for respondents whose opening bid was $2.00.

We then grouped together these SLC respondents with the subset of respondents across the 25 communities in our sample who received opening bids of $2.00 for the hypothetical CRP.11

"Yes" to whether they would be willing to pay $r_L, then we place their unknown WTP in the region ($r_L, \infty$). The likelihood function is then adjusted accordingly.

11 Ideally, in order to detect hypothetical bias, we would include RP data from communities other than Salt Lake City. By incorporating communities with fees other than $2, we would improve our ability to detect hypothetical bias in two ways. First, variation in recycling fees across communities would allow us to identify the scale parameter ($1/\sigma$) in our likelihood function and thus directly estimate the magnitude of the hypothetical bias. Second, we would be able to make more precise statements about the extent of hypothetical bias at bid levels other than $2.
With this sub-sample, we were able to estimate a joint RP-SP model, resulting in a simple probit model for the binary choice of whether or not to participate in the CRP if the fee is set at $2.00 per month. Our general hypothesis for this joint RP-SP model is:

**H1**: Those respondents bidding on the hypothetical program will have a significant response effect, resulting in an upward effect on estimated WTP relative to the SLC respondents.

In other words, we expect respondents who were asked to value the hypothetical program will, all else equal, state a higher WTP for that program than those who were asked to value their existing program. This would be evidence of hypothetical bias in our data, based on a direct comparison of SP and RP responses.

*Mitigating Hypothetical Bias*

As mentioned in the Introduction, we test two innovations in survey design to mitigate hypothetical bias—cheap talk and preference uncertainty. Cheap talk refers to a statement read immediately prior to the WTP questions that reminds the respondent to carefully consider her maximum WTP given that she will not actually have to pay for the good or service. Our cheap talk statement reads

... studies have shown that many people say they are willing to pay more for curbside recycling than they actually will pay when (it/curbside recycling) becomes available in their community. For this reason, as I read the next two curbside recycling fees, please imagine your household *actually* paying them.

This statement is noticeably shorter than the scripts used by Cummings and Taylor (1999) and List (2001). By including a binary variable in the $X_i$ vector of equation (1) for those that randomly received cheap talk (CHEAP TALK), we will test whether our shortened cheap-talk script is an effective means for reducing the upward bias often found in contingent valuation. Our general hypothesis is:

Since we are only employing SP and RP data exclusively at the $2$ level, the estimates of hypothetical bias are only suggestive of the degree of hypothetical bias at other bid levels. Unfortunately, we are precluded from using RP data from communities other than Salt Lake City because our data set contains responses from only three other communities with voluntary CRPs. Two of the three are relatively small communities and as a result of proportionate sampling, we have insufficient observations to use in the econometric analysis. In the third community, the voluntary CRP is substantially different from the one described in the hypothetical program and therefore the responses are not directly comparable to the SP responses.
**H2**: CHEAP TALK will have a significant response effect, resulting in a downward effect on estimated WTP.

Those households that did not receive cheap talk were given the following certainty question immediately after their valuation questions: "How sure are you of the answer you just gave to the previous question?" The options were very sure (VERY SURE), somewhat sure (SOMEWHAT SURE), or not sure (NOT SURE). Our general hypothesis is:

**H3**: Controlling for the presence of uncertainty will have a significant response effect, resulting in a downward effect on estimated WTP.

As is discussed in further detail below, we test this hypothesis using two different approaches. The first is similar to our test of the CHEAP TALK variable – we simply include a series of binary variables in the $X_i$ vector of equation (1) reflecting the respondent's answer to the certainty question. The second approach is a re-coding of the data, where final WTP responses to the DBDC questions of those respondents answering NOT SURE are placed in the interval $(0, \tau L)$. This approach is closest to that used by Blumenschein et al. (1998), where a respondent who answers "yes" to the question of would she (hypothetically) purchase a pair of sunglasses costing $X is only counted as a "yes" if she is also "definitely sure" of her response. Otherwise, the respondent is recorded as having answered "no". Blumenschein et al. refer to this recoding as "a conservative interpretation of the dichotomous choice contingent valuation method."

**RECYCLING SURVEYS AND DATA**

We administered two surveys for this study. The first – a household survey – was administered over the phone by the Oregon Survey Research Laboratory (OSRL) to approximately 1000 households located in 35 different communities throughout Utah with population sizes greater than 1000 residents. The communities were selected through proportionate random sampling in order to reflect a rough one-to-three split in the population between those communities with and without curbside recycling, respectively. Households were then randomly sampled within each community to ensure a roughly equal number of households.

12 Copies of the survey instruments are available on-line at http://cc.usu.edu/~aadland/research/recycling/.
with and without curbside recycling. OSRL reports that the average survey took approximately seven minutes to conduct, and response rates were approximately 75%.\textsuperscript{14} The second survey – a recycling coordinators survey – was also administered over the phone to each of the recycling coordinators in the 35 communities. This survey provided background information on each of the communities, as well as verification of the households’ responses.

We find it convenient to partition the household survey questions into five categories. The first category includes questions regarding a household's motivation for recycling. The variables are (1) “whether you feel an ethical duty to recycle to help the environment” (ETHICS), (2) “whether you believe recycling saves you money by either directly turning in recyclables or by using a smaller garbage container” (MONETARY), and (3) “whether an ethical duty to help the environment is the primary reason why you recycle” (PRIMARILY ETHICS). The inclusion of these variables enables us to test whether or not different motivations for recycling produce systematically higher WTP values, all else equal.

The second category includes a series of questions related to the status of dropoff recycling in the community and the household’s usage of dropoff recycling. Since dropoff recycling can be considered a substitute for curbside recycling, we will test whether certain characteristics of the dropoff-recycling program, such as presence or usage of dropoff facilities influence a household’s WTP for curbside service. The variables in this category include (1) “whether or not there are dropoff recycling facilities in the community” (DROP OFF) and (2) “whether the household is a frequent user of the dropoff facilities” (DROP OFF USER).

Demographic variables such as age, gender, household, membership in environmental organizations, household income level, and education levels comprise the third category of variables. Age, income, and education levels are broken down into binary variables as follows: \( 18 \leq \text{Age} \leq 35 \) (YOUNG); \( \text{Age} \geq 65 \) (OLD); \( \$25,000 \leq \text{Income} \leq \$50,000 \) (MEDIUM INCOME); \( \text{Income} \geq \$50,000 \) (HIGH INCOME); and whether the highest degree earned is either a high school diploma or GED (HIGH SCHOOL), associates degree (ASSOCIATE), bachelors degree (BACHELORS), or masters/doctoral/professional degree (GRADUATE).

The fourth category includes variables that either elicit household valuation of recycling or refine the manner in which the values are elicited. As mentioned previously, a DBDC format is

\textsuperscript{13} OSRL uses random-digit-dialing to select households into the sample, and the computer-aided-telephone-interviewing system (CATI) to interview households and record their responses (http://darkwing.uoregon.edu/~osrl/).
used, whereby respondents who do not have curbside recycling available in their communities (or are unaware that they do) are presented with a hypothetical CRP, and then offered a randomly chosen opening bid value followed by additional bids scaled to the opening bid. Since respondents who have curbside recycling available (whether or not they actually use it) are not presented with a hypothetical CRP, their opening bids are replaced by what they believe is the monthly fee for curbside service. Additional bids are then scaled to this opening amount.

The fifth and final category is only included for households with an existing CRP (whether or not they actually use it). Variables in this category include (1) community-specific binary variables, (2) whether the program is mandatory (pay for the service regardless of whether it is used) or voluntary, and (3) differences between the actual and perceived fee for curbside recycling (FEE DIFFERENTIAL). The community-specific binary variables control for unobserved heterogeneity across different types of communities. The FEE DIFFERENTIAL variable captures the household's awareness of the cost associated with participating in the CRP.

Summary statistics for the variables mentioned above are shown in Table 1. First, notice that respondents appear to recycle out of both an ethical duty to help the environment (90%) and to save money (80%), with an ethical duty being the primary motivation (68%). Second, for those respondents given the preference-certainty question, most said they were “very sure” (47%) about their responses to the valuation questions. Third, people residing in communities with CRPs who stated that they knew the curbside recycling fee (45%), tended to overstate the amount by an average of about $3.16 per month. Fourth, a little over one third (38%) of the respondents who live in a community with dropoff facilities consider themselves frequent dropoff users.

[Insert Table 1]

In terms of demographics, our sample appears fairly representative of the Utah population, except for the fact that we over sample women. Since women tend to be willing to pay more for curbside recycling, all else equal, the WTP of the “typical” household is ultimately evaluated at the population average, which is approximately a 50-50 split between men and women.

---

14 Our response rate compares favorably with Whitehead et al.’s (2000) response rate of 71%. These numbers reflect the definitions typically used by survey research laboratories to calculate response rates. The rates would change based on the definition used. We thank Paul Jakus for this insight.

15 Ultimately, only households stating values between $2 and $10, which conform to the range of opening bids for the hypothetical CRP, are included in the econometric analysis.
Imputing this population average for the typical household ultimately reduces our mean WTP estimate by approximately 15 cents.

**ECONOMETRIC RESULTS**

Referring to Table 2, note that we estimate (3) using two separate models. The first model is used to examine the effects of preference uncertainty as an explanatory variable (NOT SURE is the omitted binary variable). The second model is used to examine the effects of cheap talk (those not receiving cheap talk produce the omitted binary variable). Two separate models are necessary because in the survey instrument design, only those that did not randomly receive cheap talk were given preference-certainty questions. As such, it is not possible to simultaneously have the omitted category be those not receiving cheap talk and those stating “unsure” to the preference-certainty questions.

[Insert Table 2]

Focusing on either model in Table 2, notice first that the mean estimated WTP across our entire sample is approximately $7.58 per month. The estimated distribution for WTP within our sample is shown in Figure 1. The distribution appears to be fairly symmetric and thus is consistent with the similarity of the mean WTP and median WTP ($7.65 per month) estimates. These estimates are, however, considerably higher than those reported in Lake et al. (1996), Tiller et al. (1997), and Aadland and Caplan (1999), and are higher than the current fee of all but one of the communities in our sample that are reporting inadequate earnings from their CRPs.\(^{16}\)

[Insert Figure 1]

Second, we find several individual and community-specific characteristics that are significantly related to WTP for curbside recycling. Those who are willing to pay the most for recycling are (1) young; (2) female; (3) highly educated; (4) motivated to recycle because of either a monetary reward (e.g., smaller garbage container) or an ethical duty to help the

---

\(^{16}\) To our knowledge, these are the only other studies of household valuation of curbside recycling. Aadland and Caplan (1999) report a mean WTP for curbside recycling of approximately $2.05 per month per household. Tiller et al. (1997) report a mean WTP for dropoff recycling of approximately $4.00 per month per household. Lake et al.
environment; (5) members of an environmental organization; (6) residing in a large household; (7) not frequently using dropoff-recycling facilities; and (8) overstating the current fee for their curbside service.

Several of these effects (age, gender and education effects) are similar to those found in Aadland and Caplan (1999).\(^{17}\) However, Aadland and Caplan did not include effects (4) - (8) in their empirical model. Effects (4) and (5) are as expected, however to our knowledge (6) - (8) have not previously been tested and are therefore worthy of note. We had no prior expectation on (6), and therefore entertain several possible hypotheses regarding its sign. On the one hand, a larger household may experience larger costs associated with organizing the recycling task among its members, and thus is willing to pay less, all else equal. On the other, if the household derives passive-use value from recycling, then its payoff will be larger due to its corresponding generation of more recyclable materials. In this case, our results suggest that the latter effect outweighs the former.

One explanation for effect (7) is that households presently using dropoff facilities, all else equal, may be revealing a relatively low travel cost associated with dropoff recycling. The marginal value associated with curbside pickup (due to its added convenience) is therefore likely to be lower for these households. The effect in (8) implies a positive correlation between expected fees and WTP values (i.e., the more an individual thinks she would be paying if she were participating, the higher is her maximum WTP).

The estimation results for the RP-SP model described above in hypothesis H1 are presented in Table 3. The important variable is the dummy variable SP, which equals one for those valuing the hypothetical CRP and zero otherwise. The SP variable is positive and statistically significant, indicating an upward hypothetical bias in the SP data (i.e., we fail to reject H1) after controlling for differences in age, income level, education, household size and residence in urban versus rural communities. However, as in standard probit models, where the threshold value is not randomized, the coefficients are only identifiable up to a scale factor involving the standard deviation of the error term. The marginal effect on the probability of the true WTP being greater than $2.00 is, however, identified. As shown in Table 3, this estimate equals 0.178. In other

\(^{17}\) Aadland and Caplan (1999) also found evidence of income effects. The income coefficients in this study have the expected sign but are marginally insignificant.
words, values reported from CVM are approximately 18% more likely to exceed the $2.00 fee than the values elicited from those revealing their preferences through actual market outcomes.

[Insert Table 3]

As mentioned above, we use two separate methods to test the relationship between stated WTP and preference uncertainty. We begin by simply introducing the VERY SURE and SOMEWHAT SURE binary variables directly into the WTP equation (Model #1 in Table 2). Using this method, we find that individuals who were randomly given the certainty questions and claimed to be "very sure" regarding their stated WTP, tended to state amounts that were approximately $1.60 less, all else equal, than those who were unsure of their responses. This suggests that individuals who are less enthusiastic about recycling, or at least are not willing to pay as much for the service, are the most adamant about their feelings. Based on this approach to controlling for preference uncertainty, we therefore cannot reject hypothesis H3.

As a further test of preference-certainty effects, we re-coded the valuation responses in a manner similar to Champ, et al. (1997), Blumenschein et al. (1998) and Berrens et al. (2000). In our re-coding scheme, we assign WTP* to the (0, τ₀) region for all individuals who answered that they were "unsure" of their responses to the valuation questions. Using this approach, the estimated mean WTP for those who received the preference-certainty questions falls $0.50 from $7.42 to $6.92 per month. These conservative welfare estimates using the re-coded data may represent more appropriate values for risk-averse policymakers who are concerned about overstating projected revenues.

Referring to Model #2 (in Table 2), we find that our shortened version of Cummings and Taylor's (1999) "cheap talk" is, in the aggregate, rather ineffective in mitigating hypothetical bias. While the coefficient on the cheap-talk binary variable is negative, it is small in magnitude (-$0.22) and is statistically insignificant at the 10% level. However, it seems reasonable that cheap talk may be at least marginally more successful for certain types of individuals.¹⁸

To investigate this conjecture, we interacted the cheap-talk binary variable with several individual characteristics. The results from this exercise are shown in Table 4. First, notice that

¹⁸ Our motivation for this hypothesis is List's (2001) similar finding for dealer vs. non-dealer bids in hypothetical auctions for a mint baseball card. List finds that cheap talk (using a long script similar to Cummings and Taylor's) fails to mitigate hypothetical bias in dealer valuation exercises, but is effective in controlling bias in auctions involving ordinary consumers.
the coefficients on the single interactive terms are generally negative, larger in magnitude than
the coefficient on the non-interactive term, although not always statistically significant. The
exceptions are the coefficients on the environmental-organization and medium-income variables,
which indicate that members of an environmental group and households with medium incomes
who received cheap talk state significantly lower WTP values than their respective counterparts
who did not receive cheap talk. The other exceptions are the positive (albeit statistically
insignificant) coefficients on the interactive terms with GRADUATE and HIGH INCOME.

A further refinement, as shown by the double interaction terms in Table 4, however, indicates
that certain subgroups (e.g., young females and females residing in households with medium
incomes) do report significantly lower WTP values than their respective counterparts who do not
receive cheap talk. Therefore, when taken as a whole, we find mixed evidence regarding
hypothesis H2 (i.e., whether cheap talk is effective in reducing hypothetical bias in the stated WTP
values for curbside recycling). Similar to List's (2001) finding for private-good auctions,
certain types of individuals appear to respond to cheap talk, while others do not.

[Insert Table 4]

Turning now to goodness of fit, a common measure in limited dependent variable models is
McFadden’s pseudo $R^2$, which is one less the ratio of the unrestricted to the restricted log
likelihood value. McFadden’s pseudo $R^2$ for the model is approximately 0.08. While this does
not appear to be indicative of a good fit, recall that the probit model does not maximize any
specific measure of fit, as in the linear regression model. Also, there is a large literature
acknowledging the difficulty of interpreting pseudo $R^2$ measures between zero and one (see for
example Amemiya, 1981 and Greene, 2000). In response, we calculate two additional goodness-
of-fit measures.

First, the likelihood ratio statistic (used to test the hypothesis that all the coefficients – other
than the intercept – are jointly equal to zero) is 202.3, with a 5% critical chi-square value of 49.8.
Therefore, the explanatory variables taken as a whole are able to explain a significant amount of
the variation in the dependent variable. Second, we calculate a matrix of correct and incorrect

---

19 This discussion refers to Model #1. The results from Model #2 are very similar.
predictions in Table 5. Summing the elements along the main diagonal shows that we were able to correctly place 41% of the individuals in the correct interval. If we further include elements along the secondary diagonal (i.e., also consider relatively close misses), our prediction percentage rises to 86%. When taken together, we consider these measures to be suggestive of a relatively good overall fit for our joint model.

[Insert Table 5]

CONCLUSIONS AND POLICY IMPLICATIONS

Based on a sample of over 1000 households from Utah (including households located in communities with mandatory and voluntary CRPs, as well as households in communities without CRPs), we find that the mean WTP for curbside recycling across our entire sample is approximately $7.58 per month. Young, well-educated women who are members of environmental organizations, who recycle out of an ethical responsibility for the environment, who are not frequent dropoff users, and reside in large households are willing to pay the most for the curbside service. Further, using data combined from stated and revealed preferences, we find statistically significant hypothetical bias in the SP data. We are able to partially mitigate this bias, by introducing “cheap talk” and controlling for “preference uncertainty”. By setting CHEAP TALK equal to one, re-coding the WTP data using the preference-uncertainty information, and adjusting for an over sampling of females, the predicted WTP then falls by $0.87 per month.

This information should be useful to both researchers interested in valuing goods with passive-use value and policymakers considering allocating resources toward curbside recycling. Researchers can use the detection and mitigation methods explored in this study to appropriately adjust the estimated non-market values for similar environmental goods. Furthermore, local policymakers can use our analysis to estimate a revenue function that relates projected revenues to recycling fees. In Figure 2, we present a graph relating projected revenues from a voluntary CRP to a range of fees. The projected revenues are calculated using our estimates from the entire sample under the assumption that households will participate if their predicted WTP is

---

20 The total number of observations sums to 816 rather than 876 because, as mentioned earlier, some respondents answered “Don’t Know” to some of the valuation questions. For these 60 households, their unknown WTP does not fit nicely into one of the five categories, but instead overlaps one or more of the intervals.
greater than the fee. The product of the fee and the number of participating households then gives the total projected revenue.

[Insert Figure 2]

Coupled with data on the projected costs of providing a curbside recycling service in their community, policymakers could then equate the incremental costs of providing the service to the incremental revenue associated with providing the service to additional individuals (through the relationship between fee and participation changes) to determine the efficient allocation of resources toward curbside recycling.

Future research on these issues might proceed along two lines. First, a broader study of household recycling behavior could be undertaken in order to test the robustness of the welfare estimates presented in this paper. Ideally, this would be a larger regional or national study, thus expanding the scope of both households and communities. Second, the effects of cheap talk and preference uncertainty in mitigating hypothetical bias deserves further testing. Cheap talk was found to be mildly effective in mitigating bias for certain types of respondents – those who are perhaps more prone than average to overstate their WTP values in the first place. We admittedly used a shortened cheap-talk script, but one that is more appropriate for telephone interviews. Future studies might vary the script length in an effort to estimate the marginal effect of script length in mitigating hypothetical bias. Future research might also expand upon the preference-uncertainty approaches used in this study to test their effectiveness in a variety of settings and formats.
REFERENCES


Table 1. Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethical Duty</td>
<td>0.8978</td>
<td>0.3031</td>
<td>0.0000</td>
<td>1.0000</td>
<td>871</td>
</tr>
<tr>
<td>Monetary</td>
<td>0.7971</td>
<td>0.4024</td>
<td>0.0000</td>
<td>1.0000</td>
<td>818</td>
</tr>
<tr>
<td>Ethics First</td>
<td>0.6844</td>
<td>0.4652</td>
<td>0.0000</td>
<td>1.0000</td>
<td>583</td>
</tr>
<tr>
<td>Very Sure</td>
<td>0.4747</td>
<td>0.5000</td>
<td>0.0000</td>
<td>1.0000</td>
<td>375</td>
</tr>
<tr>
<td>Somewhat Sure</td>
<td>0.3200</td>
<td>0.4671</td>
<td>0.0000</td>
<td>1.0000</td>
<td>375</td>
</tr>
<tr>
<td>Fee Differential</td>
<td>-3.1642</td>
<td>2.9865</td>
<td>-8.0200</td>
<td>4.0000</td>
<td>237</td>
</tr>
<tr>
<td>Fee Known</td>
<td>0.4473</td>
<td>0.4983</td>
<td>0.0000</td>
<td>1.0000</td>
<td>237</td>
</tr>
<tr>
<td>Dropoff</td>
<td>0.8151</td>
<td>0.3885</td>
<td>0.0000</td>
<td>1.0000</td>
<td>795</td>
</tr>
<tr>
<td>Dropoff User</td>
<td>0.3787</td>
<td>0.4854</td>
<td>0.0000</td>
<td>1.0000</td>
<td>647</td>
</tr>
<tr>
<td>Young</td>
<td>0.3196</td>
<td>0.4666</td>
<td>0.0000</td>
<td>1.0000</td>
<td>876</td>
</tr>
<tr>
<td>Old</td>
<td>0.1450</td>
<td>0.3523</td>
<td>0.0000</td>
<td>1.0000</td>
<td>876</td>
</tr>
<tr>
<td>Male</td>
<td>0.3619</td>
<td>0.4808</td>
<td>0.0000</td>
<td>1.0000</td>
<td>876</td>
</tr>
<tr>
<td>High School</td>
<td>0.2592</td>
<td>0.4384</td>
<td>0.0000</td>
<td>1.0000</td>
<td>872</td>
</tr>
<tr>
<td>Associates</td>
<td>0.3704</td>
<td>0.4832</td>
<td>0.0000</td>
<td>1.0000</td>
<td>872</td>
</tr>
<tr>
<td>Bachelors</td>
<td>0.2076</td>
<td>0.4058</td>
<td>0.0000</td>
<td>1.0000</td>
<td>872</td>
</tr>
<tr>
<td>Graduate</td>
<td>0.0940</td>
<td>0.2920</td>
<td>0.0000</td>
<td>1.0000</td>
<td>872</td>
</tr>
<tr>
<td>Household Size</td>
<td>3.2255</td>
<td>1.8271</td>
<td>1.0000</td>
<td>13.000</td>
<td>869</td>
</tr>
<tr>
<td>Environ. Org.</td>
<td>0.0708</td>
<td>0.2566</td>
<td>0.0000</td>
<td>1.0000</td>
<td>876</td>
</tr>
<tr>
<td>Med Income</td>
<td>0.4203</td>
<td>0.4939</td>
<td>0.0000</td>
<td>1.0000</td>
<td>759</td>
</tr>
<tr>
<td>High Income</td>
<td>0.3953</td>
<td>0.4892</td>
<td>0.0000</td>
<td>1.0000</td>
<td>759</td>
</tr>
</tbody>
</table>

Notes. The varying sample sizes reflect the elimination of "Don’t Know" responses, refusal to answer, or households that did not receive the survey question.
Table 2. Estimation Results for the DBDC Model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model #1 (n=876)</th>
<th></th>
<th>Model #2 (n=876)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient</td>
<td>P-Value</td>
<td>Coefficient</td>
<td>P-Value</td>
</tr>
<tr>
<td>Ethical Duty</td>
<td>4.389***</td>
<td>(0.000)</td>
<td>4.199***</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Monetary</td>
<td>2.035**</td>
<td>(0.045)</td>
<td>1.851*</td>
<td>(0.062)</td>
</tr>
<tr>
<td>Primarily Ethics</td>
<td>1.389***</td>
<td>(0.004)</td>
<td>1.351***</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Dropoff</td>
<td>0.331</td>
<td>(0.277)</td>
<td>0.341</td>
<td>(0.272)</td>
</tr>
<tr>
<td>Dropoff User</td>
<td>-0.735*</td>
<td>(0.062)</td>
<td>-0.775*</td>
<td>(0.052)</td>
</tr>
<tr>
<td>Young</td>
<td>2.734***</td>
<td>(0.000)</td>
<td>2.938***</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Old</td>
<td>-1.854***</td>
<td>(0.002)</td>
<td>-1.757***</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Male</td>
<td>-1.070***</td>
<td>(0.005)</td>
<td>-0.985***</td>
<td>(0.009)</td>
</tr>
<tr>
<td>High School</td>
<td>0.501</td>
<td>(0.281)</td>
<td>0.361</td>
<td>(0.340)</td>
</tr>
<tr>
<td>Associates</td>
<td>0.672</td>
<td>(0.215)</td>
<td>0.430</td>
<td>(0.308)</td>
</tr>
<tr>
<td>Bachelors</td>
<td>1.457*</td>
<td>(0.057)</td>
<td>1.244*</td>
<td>(0.090)</td>
</tr>
<tr>
<td>Graduate</td>
<td>3.440***</td>
<td>(0.001)</td>
<td>3.225***</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Household Size</td>
<td>0.243**</td>
<td>(0.017)</td>
<td>0.259**</td>
<td>(0.012)</td>
</tr>
<tr>
<td>Environ. Org.</td>
<td>1.366**</td>
<td>(0.042)</td>
<td>1.309**</td>
<td>(0.049)</td>
</tr>
<tr>
<td>Med Income</td>
<td>0.629</td>
<td>(0.155)</td>
<td>0.662</td>
<td>(0.143)</td>
</tr>
<tr>
<td>High Income</td>
<td>0.817</td>
<td>(0.110)</td>
<td>0.793</td>
<td>(0.118)</td>
</tr>
<tr>
<td>Cheap Talk</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very Sure</td>
<td>-1.634**</td>
<td>(0.016)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somewhat Sure</td>
<td>0.810</td>
<td>(0.162)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fee Differential</td>
<td>-0.319*</td>
<td>(0.061)</td>
<td>-0.320*</td>
<td>(0.063)</td>
</tr>
<tr>
<td>Fee Understated</td>
<td>-1.627</td>
<td>(0.101)</td>
<td>-1.637</td>
<td>(0.101)</td>
</tr>
<tr>
<td>Fee Known</td>
<td>-0.796</td>
<td>(0.228)</td>
<td>-0.775</td>
<td>(0.236)</td>
</tr>
<tr>
<td>Bid</td>
<td>0.195***</td>
<td>(0.000)</td>
<td>0.193***</td>
<td>(0.000)</td>
</tr>
</tbody>
</table>

Mean WTP $7.58 / month $7.58 / month

Notes. (***) (**), and (*) refer to statistical significance at the 1, 5 and 10 percent levels respectively. The results for variables such as the constant, "Don't Know" dummy variables, and community-specific effects are not shown. The coefficient on the bid is (1/α) and all other reported coefficients reflect the actual βs with the scale factor removed. The coefficient on the Cheap Talk variable is omitted from Model #1 to avoid confusion, although it is important to recognize that it was included as an explanatory variable. The likelihood ratio statistics for Models #1 and #2 are 202.31 and 188.49, respectively. McFadden's pseudo R^2's are 0.078 and 0.073, respectively.
### Table 3. Probit Estimation Results for the Joint SP-RP Model (n= 145)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.950**</td>
<td>(0.047)</td>
</tr>
<tr>
<td>Young</td>
<td>1.349***</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Old</td>
<td>-0.023</td>
<td>(0.476)</td>
</tr>
<tr>
<td>Male</td>
<td>-0.177</td>
<td>(0.248)</td>
</tr>
<tr>
<td>High School</td>
<td>0.104</td>
<td>(0.421)</td>
</tr>
<tr>
<td>Associates</td>
<td>0.254</td>
<td>(0.307)</td>
</tr>
<tr>
<td>Bachelors</td>
<td>0.598</td>
<td>(0.128)</td>
</tr>
<tr>
<td>Graduate</td>
<td>1.145**</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Household Size</td>
<td>0.001</td>
<td>(0.494)</td>
</tr>
<tr>
<td>Med Income</td>
<td>0.076</td>
<td>(0.409)</td>
</tr>
<tr>
<td>High Income</td>
<td>0.557*</td>
<td>(0.052)</td>
</tr>
<tr>
<td>Urban</td>
<td>0.601**</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Stated Preference (SP)</td>
<td>0.475*</td>
<td>(0.054)</td>
</tr>
</tbody>
</table>

**Marginal Effect**

\[
\text{Prob(WTP}^* > \$2 \mid \text{SP}) - \text{Prob(WTP}^* > \$2 \mid \text{RP}) = 0.178
\]

**Notes.** (***)**, (**), and (*) refer to statistical significance at the 1, 5 and 10 percent levels respectively. The Urban dummy variable equals one if the respondent resides in a community with a population greater than 100,000 and zero otherwise.
Table 4. Interactive Cheap Talk Estimated Coefficients

<table>
<thead>
<tr>
<th></th>
<th>Single Interaction Variables</th>
<th></th>
<th>Double Interaction Variables</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>None Primarily Ethic Young Female Bachelors Graduate Environ. Dropoff Org. Med User Income High Income</td>
<td>None Young * Female Primarily Ethic</td>
<td>Young * Med Income Female * Primarily Ethics</td>
<td>Female * Med Income</td>
</tr>
<tr>
<td>Coefficients</td>
<td>-0.218 -0.786 -0.946 -0.218 -0.895 0.759 -3.634** -0.919 -1.234** 0.188</td>
<td>-0.218 -1.544* -1.743 -2.513** -1.131* -1.742**</td>
<td>Coefficients P Values 0.320 0.073 0.103 0.019 0.099 0.033</td>
<td></td>
</tr>
<tr>
<td>P Values</td>
<td>0.320 0.132 0.138 0.355 0.192 0.332 0.037 0.138 0.049 0.395</td>
<td>0.320 0.073 0.103 0.019 0.099 0.033</td>
<td>P Values</td>
<td>0.320 0.073 0.103 0.019 0.099 0.033</td>
</tr>
</tbody>
</table>

Notes. (***) (**), and (*) refer to statistical significance at the 1, 5 and 10 percent levels respectively. The model is the same as Model #2 from Table 2 except the cheap-talk binary variable is replaced with interactive variables. Each of the above variables are interacted with the cheap-talk binary variable. The respective omitted categories are the variables interacted with those that did not get cheap talk.
Table 5. Number of Correct and Incorrect WTP Predictions

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Actual</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(-\infty, 0)</td>
<td>(0, \tau_L)</td>
</tr>
<tr>
<td>(-\infty, 0)</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>(0, \tau_L)</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>(\tau_L, \tau)</td>
<td>13</td>
<td>60</td>
</tr>
<tr>
<td>(\tau, \tau_H)</td>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>(\tau_H, \infty)</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Totals</td>
<td>35</td>
<td>127</td>
</tr>
</tbody>
</table>

Notes. Results from Model #1 in Table 2.
Figure 1. Frequency Distribution for Predicted Willingness to Pay
Figure 2. CRP Revenue Function per 100,000 Households

Annual Revenue (in Millions of Dollars) vs. CRP Fee
Detection and Mitigation of Hypothetical Bias in Contingent Valuation
With an Application to Curbside Recycling

David Aadland and Arthur J. Caplan*

February 13, 2001

Abstract. In this paper, we use a unique curbside-recycling data set to test the effectiveness of "cheap talk" and "preference uncertainty" in mitigating hypothetical bias in contingent valuation. The sample includes two types of households – those located in communities with curbside recycling programs (mandatory or voluntary) and those in communities without curbside recycling. Using stated and revealed preference data, we detect significant hypothetical bias. Cheap talk and preference-uncertainty controls are partially effective in mitigating the bias.

JEL Classification: C35, D12

*The authors are assistant professors in the Department of Economics, Utah State University, 3530 Old Main Hill, Logan, UT 84322. We thank Patricia Gwartney (director), Emery Smith (programmer), and other employees of the University of Oregon's Survey Research Laboratory for conducting the survey for this study. We also thank Paul Jakus, John Keith, John Whitehead, and Therese Grijalva for comments on an earlier version of the paper. Ryan Bosworth provided graduate research assistance. The USU Research Initiative Program provided research funds.
INTRODUCTION

There is an active debate regarding the reliability of information from contingent valuation (CV), and more generally, regarding the issue of hypothetical bias from stated-preference methods (Diamond and Hausman, 1994; Carson et al., 2000; Arrow et al., 1993). Although there is disagreement about the extent of the bias, nearly all researchers agree that it exists. Hypothetical bias arises when people are asked to hypothetically state or select a maximum amount they are willing to pay for a good or service, even though they know they will never have to actually pay for it. Researchers have typically used two general methods to detect and mitigate this bias. The first involves innovations in the design of the survey instrument. The second involves combining results of actual market outcomes (i.e., revealed preference (RP) data) with stated preference (SP) data to improve the reliability of the resulting welfare estimates.

Recent innovations in survey design to mitigate hypothetical bias include: (1) reminder statements of substitutes, budget constraints, or the hypothetical nature of the good in question (Neill et al., 1994; Neill, 1995; Loomis et al., 1996; Cummings and Taylor, 1999; List, 2001), (2) use of referendum formats (Mitchell and Carson, 1989; Loomis, 1990; Arrow et al., 1993; Alberini, 1995b; Boyle et al., 1996; Bohara et al., 1998; Welsch and Poe, 1998; Carson et al., 2000), and (3) follow-up questioning (Cameron and Quiggin, 1994; Alberini, 1995b; Kanninen, 1995; Li and Mattsson, 1995; Blumenschein et al., 1998; Berrens et al., 2000; Champ et al., 1997). In this paper, we test two of these innovations.

The first is the use of "cheap talk" in CV, as introduced by Cummings and Taylor (1999).1 Cheap talk is information provided prior to the willingness-to-pay (WTP) questions reminding respondents that they are being asked to value a hypothetical program. Essentially, it is a scripted explanation of how the program's hypothetical nature can bias WTP responses. Cummings and Taylor find some evidence that households receiving cheap talk will, on average, report lower WTP values than those that do not (i.e., that cheap talk may correct for upward bias in WTP responses), but they also suggest the need for further empirical testing. In particular,

---

1 Although Cummings and Taylor (1999) were the first to use the game-theoretic term “cheap talk” in the context of CV, Loomis et al. (1994) employed a similar approach in an earlier study. In a laboratory experiment, Loomis et al. were able to reduce the discrepancy between hypothetical and actual WTP by issuing a reminder to their subjects (prior to the WTP question) that “...although the question is hypothetical, we want you to answer as if it were real...”