Cheap Talk Reconsidered: New Evidence from CVM

David Aadland
Department of Economics and Finance
University of Wyoming
Laramie, WY 82071-3985
Phone: 307-766-4931
Fax: 307-766-5090
email: aadland@uwyo.edu

Arthur J. Caplan
Department of Economics
Utah State University
3530 Old Main Hill
Logan, UT 84322-3530
Phone: 435-797-0775
Fax: 435-797-2701
email: acaplan@econ.usu.edu

September 13, 2012

Cheap Talk Reconsidered: New Evidence from CVM*

Abstract. Two recent studies have shown that "cheap talk" is an effective means of eliminating positive hypothetical bias in experimental and field-auction settings. We further investigate the ability of cheap talk to mitigate positive hypothetical bias in a contingent-valuation phone survey administered to over 4,000 households. Positive hypothetical bias is detected in our data by contrasting revealed and stated preference information. However, a short, neutral cheap-talk script appears to exacerbate rather than mitigate the bias. Based on this and mixed evidence from earlier studies, we suggest caution in using cheap talk as an *ex ante* control for hypothetical bias.

JEL Classification: Q26, C35

Keywords: cheap talk, contingent valuation, hypothetical bias

*The authors thank John Tarnei, Kent Miller and other employees of the Washington State University's Survey Research Laboratory for conducting the survey for this study. We also thank David Grether (coeditor), two anonymous referees, Robert Berrens, Glenn W. Harrison, and John Loomis for insightful comments on an earlier draft of the paper. This research is supported by National Science Foundation grant #0108159.

1. Introduction

The contingent valuation method (CVM) is a widely used approach for estimating the value of goods and services when market information on equilibrium prices and quantities is either unavailable or unreliable. CVM has been employed by courts and government agencies such as the U.S. Environmental Protection Agency, the National Oceanographic and Atmospheric Administration, and the U.S. Fish and Wildlife Service to assess the benefits of policies impacting the environment and damages from environmental disasters. Researchers often estimate these values through surveys that ask individuals to place a monetary value on the hypothetical provision of a good or service. Since provision of the good and the associated payment are purely hypothetical, the reliability and validity of information obtained from CVM has been the subject of lively debate (Diamond and Hausman 1994, Hanneman 1994). Proponents of CVM have attempted to develop new methodologies that either (1) mitigate *ex ante* any hypothetical bias (i.e., bias associated with the respondent misstating her maximum willingness to pay (WTP) due to the hypothetical nature of the good and payment method), or (2) calibrate the welfare estimates *ex* post (List and Shogren 1998, Harrison et al. 1999).

Recently, Cummings and Taylor (1999) and List (2001) have advocated the use of "cheap talk" to mitigate *ex ante* the effects of hypothetical bias in CVM. In the context of CVM, cheap talk refers to explicit warnings about the problem of hypothetical bias provided prior to respondents' valuation of the good. Cummings and Taylor (CT hereafter), in a series of lab experiments with students, find that cheap talk successfully eliminates hypothetical bias in valuation responses for a variety of public goods. List tests a similar script for private goods using sportscard auctions and finds that cheap talk is effective in eliminating hypothetical bias for non-dealers, but not for dealers. The cheap-talk scripts used in both of these studies are almost identical in length and content. They each provide lengthy descriptions of *positive* hypothetical bias.

In order for cheap talk (such as that applied by CT and List) to be a useful design element in CVM surveys, the script needs to be general so that, unlike *ex post* calibration, it can be easily applied across a wide array of non-market goods without requiring *ex ante* information on the degree of hypothetical bias in the data. Unfortunately, the scripts used in CT and List are not easily generalized.[1] Both scripts refer to a baseline degree of hypothetical bias by comparing the outcomes of preliminary experiments with hypothetical and real payment mechanisms for the goods in question. In CVM research, such prior information regarding the degree of hypothetical bias is typically unavailable or too expensive to produce in the field. The researcher must therefore presume the degree of hypothetical bias that exists in the population and subsequently calibrate the specific wording of his cheap-talk script based solely on this presumption. The more unique the population or good in question, the more potentially problematic is this calibration-by-presumption approach.

Our research addresses this concern by testing a more "neutral" version of cheap talk that can easily be generalized to other goods or services. We administered a telephone survey to over 4,000 households regarding their WTP for a curbside recycling program (CRP). Unlike the pure public goods used in the CT experiments and the private good in List's field auctions, curbside recycling can be considered an impure public good because it provides both potential private benefits (in the forms of added convenience relative to dropoff recycling, reduced garbage fees, and a "warm glow" from helping the environment), and public benefits by reducing the stream of

[1]In addition to their main script, CT report similar results using a modified script that replaces the specific percentages of people in previous studies who voted "yes" for hypothetical and real referenda, with a statement indicating that "on average" more people voted "yes" for hypothetical referenda (see CT, pages 659-660). Although the modified script is more general in the sense of not reporting the magnitude of hypothetical bias, it still informs the subjects that the bias is positive.

waste going to landfills (Andreoni 1990). As part of our survey design, we randomly assign our sample into three groups; the first receiving no cheap talk, the second receiving a short-script version of cheap talk, and the third receiving cheap talk with a reminder of budget constraints and substitutes. Aside from using a shorter script (to be compatible with phone surveys), the primary difference between our cheap-talk scripts and those used by CT and List is the manner in which hypothetical bias is described to the survey respondents. While CT and List state that hypothetical bias leads people to *overstate* their true WTP, the cheap-talk scripts in our survey are purposefully neutral. Instead, they inform respondents that hypothetical bias leads people to *misstate* their true WTP.

We crafted neutral cheap-talk statements for two reasons. First, we wished to avoid criticisms that we are simply "layering" on another type of bias, one that may be independent of the observed hypothetical bias itself.[2] Second, although most research suggests that hypothetical bias is positive (e.g., Arrow et al. 1993, Hannemann 1994, and Diamond and Hausman 1994), this is not always the case (Dickie et al. 1987, Adamowicz et al. 1994, Carson et al. 1996, Nestor 1998, and Haab et al. 1999). As a result, researchers cannot be certain *ex ante* whether hypothetical bias will be positive or negative in their population, much less know its magnitude. Unlike CT and List, we had little evidence regarding the direction and magnitude of hypothetical bias for our population or the good in question. Therefore, we decided to err on the side of caution so as not to induce any additional bias. Much to our surprise, this seemingly innocuous change in verbiage from a directed to a neutral script caused respondents receiving cheap talk to state *higher* WTP than those not receiving cheap talk.

[2] This hypothesis of a "layering effect" is refuted by CT for one of their sub-groups. However, we feel that it is still an open question, particularly for CVM surveys.

This counterintuitive result is robust across models, types of households, and types of recycling programs. Moreover, the result is in sharp contrast to our earlier study of household recycling behavior in Utah, where we found that a similar short-script version of cheap talk with directed rather than neutral wording caused households to state lower WTP on average (Aadland and Caplan 2003). That study was based on the same type of good, the same assumptions concerning the household's underlying preference structure, the same empirical approach, similar sets of explanatory variables, and a sample from within the population used for this study. Because of the "procedural invariance" between the two studies (Kahneman and Tversky 1984) and as a result of the studies' distinctive findings, we are led to question the efficacy of cheap-talk statements in mitigating hypothetical bias in CVM surveys.[3]

The next section provides examples of the cheap-talk designs used in previous studies. In Section 3, we briefly describe the survey and cheap-talk scripts used in our study. Section 4 reports our empirical evidence on hypothetical bias and cheap talk. Section 5 summarizes our overall findings.

2. Previous Use of Cheap Talk

CT are the first to use the game-theoretic terminology "cheap talk" in the context of CVM. Cheap talk differs from standard reminder statements about substitutes and budget constraints in that the script explicitly warns respondents about the potential problem of hypothetical bias. Loomis et al. (1994) and Neil (1995) find that short-scripted reminder statements (without cheap talk) are ineffective in altering respondents' stated WTP for a public good in a hypothetical setting. However, CT find that a cheap-talk script openly discussing the existence of positive

[3] Several studies touch on the importance of specific wording or instrument calibration effects, as in Harrison (2002).

hypothetical bias prior to voting on public good referenda eliminates the bias in an experimental setting, in the sense that the results from the cheap-talk and actual referenda are statistically indistinguishable.[4] An excerpt from CT's cheap-talk script is given below:

> *... in a recent study, several different groups of people voted on a referendum just like the one you are about to vote on. Payment was hypothetical for these groups, as it will be for you. No one had to pay money if the referendum passed. The results of these studies were that on average 38 percent of them voted "yes". With another set of groups with similar people voting on the same referendum as you will vote on here, but where payment was real and people really did have to pay money if the referendum passed, the results on average across the groups were that 25 percent voted yes. That's quite a difference, isn't it?*
>
> *We call this a "hypothetical bias." Hypothetical bias is the difference that we continually see in the way people respond to hypothetical referenda as compared to real referenda…*

List reads a nearly identical cheap-talk script to market participants in a field auction for sportscards. He finds that cheap-talk statements are effective in eliminating hypothetical bias, but only for ordinary consumers (i.e., non-dealers). Dealers, who presumably have more market experience in valuing sportscards, are not influenced by cheap talk. Similarly, Brown et al. (2003) report that CT's cheap talk script is effective at high referendum bid levels but ineffective at low bid levels.

Poe et al. (2002) find that the following short script,

> *I have one caution though. For programs like this it's often the case that more people say they would sign up than actually do sign-up. Utilities in other parts of the country have found that eight times as many people say yes to similar programs as actually take part in them. With this in mind…*

does not have a statistically significant effect on the participation decisions of individuals valuing green power and tree planting in New York via a provision-point mechanism CVM

[4] Two other studies, Loomis et al. (1996) using a short script somewhere between a reminder statement and cheap talk and Murphy et al. (2003) using CT's cheap talk script, find that while these statements do not eliminate hypothetical bias, they do reduce it.

survey. Aadland and Caplan (2003) also employ a shorter version of cheap talk than CT and List. Similar to List, they find that the effectiveness of cheap talk varies by type of household. In particular, those households who might be expected to suffer the most from positive hypothetical bias (e.g. those motivated to recycle by an ethical duty or who are members of an environmental organization) also tend to lower their stated WTP the most in response to cheap talk. Their cheap-talk script reads,

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

To our knowledge, this is the first evidence that short cheap-talk scripts can be effective in mitigating positive hypothetical bias in CVM surveys.[5]

3. Cheap-Talk and Survey Design

During the winter of 2002, we conducted a telephone survey about recycling behavior to over 4,000 households in 40 western U.S. cities with populations over 50,000.[6] After the surveys were completed, the households were divided into two main groups. One group (henceforth "CRP-H households") includes households who either do not have a CRP in their community or who reside in a community with a CRP but are unaware of its existence. For this group, we described the following hypothetical CRP:

---

[5] Two working papers also deserve mention. First, Cummings et al. (1995) find that while "heavy" cheap talk tends to offset positive hypothetical bias, "light" cheap talk actually tends to increase the upward bias in a public good valuation experiment. Second, Bulte et al. (2003), in a field study using "netbox" technology (which enables respondents to retrieve and return questionnaires via a television set) find that a shortened version of cheap talk is effective in mitigating positive bias. We note, however, that the light and shortened scripts of Cummings et al. and Bulte et al. are still much longer than those used by Poe et al. (2002) and Aadland and Caplan (2003).

[6] The survey was administered by the survey research laboratory at Washington State University. A list of the 40 cities included in our sample and the survey instrument are available at www.uwyo.edu/aadland/research/recycle/.

At the beginning of the survey, you said that your community does not currently have a curbside recycling service. For the next few questions, please imagine that you COULD have a service that regularly collects aluminum cans, cardboard, glass, paper, plastic and tin cans. Your household (would/would not) need to sort your recyclables into separate bins and pay a fee for the recycling service, in addition to your current monthly garbage collection fee. Now we are going to ask you some questions about your household's willingness to pay for this type of curbside recycling service.

The second group (henceforth “CRP-A households”) includes households residing in communities with an actual CRP and who know of its existence. These households were asked hypothetical questions regarding their WTP for their community’s existing program, regardless of whether they actually participate in it.[7] Because they reside in communities with mandatory or voluntary CRPs, we further divided the CRP-A households into two sub-groups.[8] CRP-A households residing in communities with a voluntary CRP are assigned to the CRP-V sub-group, while those residing in communities with a mandatory CRP are assigned to the CRP-M sub-group.

We then randomly assigned all respondents (i.e., all CRP-H and CRP-A households) with equal probability to one of three cheap-talk groups. The first group received no cheap-talk statement and proceeded directly to the valuation questions. The second group received the following short cheap-talk script, which was read prior to the ensuing WTP question:

As you prepare to answer the next few questions, please keep in mind that in previous surveys we have found that the amounts that people SAY they are willing to pay for curbside recycling are sometimes different from the amounts that they would ACTUALLY be willing to pay when curbside recycling became available in their community. For this reason, as I read the following curbside recycling fees, please imagine your household is ACTUALLY paying them.

---

[7] This presumes that non-participating CRP-A households are not only aware of the CRP’s existence, but are also familiar with the program’s main attributes (e.g., such as those mentioned in the hypothetical description) via observing their neighbors who do recycle or through occasional exposure in their local media.
[8] By “voluntary” we mean that the household pays for the program only if it has signed up for it. “Mandatory” means that the household pays for the program regardless of whether it has signed up for it.

This cheap-talk script is substantially shorter than that used by CT and List (so as to be compatible with a phone survey) and intentionally does not take a stand regarding the direction of hypothetical bias. The third group of households received the following "long" cheap-talk script, which added explicit reminders to the household about budget constraints and alternatives to recycling:

> *As you prepare to answer the next few questions, please keep in mind the following three things. First, keep in mind your household budget. In a typical month, at what price would your household be able to afford curbside recycling? Second, keep in mind that there are alternatives to curbside recycling such as recycling drop-off centers and landfills. And third, keep in mind that in previous surveys we have found that the amounts that people SAY they are willing to pay for curbside recycling are sometimes different from the amounts that they would ACTUALLY be willing to pay when curbside recycling became available in their community. For this reason, as I read the following curbside recycling fees, please imagine your household is ACTUALLY paying them.*

Our decision to test variations in script length and reminders about budget constraints and substitutes is motivated by mixed results in the literature. As mentioned above, CT find that long-scripted cheap talk is effective in an experimental setting. List finds that similar long-scripted cheap talk is effective in field auctions for dealers but ineffective for non-dealers. Brown et al. find that the effectiveness of long cheap talk depends upon the bid level. Similar to List, Aadland and Caplan (2003) find that a short-scripted version of cheap talk is effective only for certain types of households. Poe et al. report that short-scripted cheap talk is ineffective in eliminating hypothetical bias. Loomis et al. (1994) and Neil find that reminders about budget constraints and substitutes are also ineffective. In sum, the evidence regarding the effectiveness of various cheap-talk and reminder statements is anything but clear.

4. Empirical Results

In this section, we report three sets of empirical results. The first set documents the existence of positive hypothetical bias in our data. That is, we find that CRP-H households are, all else equal, more likely to (state that they will) participate in a CRP than the CRP-V households.[9] The second set of results documents the unconditional effects of cheap-talk scripts on household yes/no responses to randomized opening bid values (which effectively determine their hypothetical participation rates). Here, we simply contrast the responses of all households who received a cheap-talk script with those households that were not subjected to cheap talk.[10] The advantage of examining the unconditional participation rates is that they are transparent and are not dependent on any particular econometric model. Finally, our third set of results provides conditional evidence on the effects of our cheap-talk scripts. We report the coefficient estimates associated with cheap-talk dummy variables from an econometric model where we control for a plethora of potential differences across treatments, groups, and household demographics (see the Appendix for definitions of the control variables).

4.1 Estimates of Hypothetical Bias

We begin by documenting the existence of positive hypothetical bias in our CVM responses. Toward this end, we compare the (hypothetical) participation decisions of CRP-H households with the (actual) participation decisions of CRP-V households. Our survey was designed to facilitate such a comparison by choosing appropriate opening bid values and describing

[9] Only CRP-V households are able to reveal their true preferences and thus provide a benchmark for determining the degree of hypothetical bias in the CRP-H households' responses.

[10] Recall that all households in our sample received a hypothetical WTP question; for CRP-H households the WTP question pertained to a hypothetical CRP while for CRP-A households the question referred to their existing CRP.

hypothetical CRPs that are similar to existing voluntary programs in our sample.[11] To detect hypothetical bias we pool CRP-V households and CRP-H households who have valued hypothetical CRPs with attributes similar to the existing voluntary programs valued by CRP-V households.[12] We then estimate a single-bounded probit model (with known thresholds) for the decision of whether to participate in a CRP, while controlling for a host of demographic, program, and community attributes such as age, education, income, employment status, motivation for recycling, degree of sorting required for CRP, availability and use of substitutes to recycling, and so on.

Our null hypothesis of no hypothetical bias is tested by observing the statistical significance of the coefficient on the dummy variable for whether the participation decision is hypothetical (i.e., based on the CRP-H household's valuation of the hypothetical program) or real (i.e., based on the CRP-V household's actual decision of whether to participate in its community's CRP).[13] If we find that there is no statistical difference between the participation rates of these two groups of households then, all else equal, we conclude that hypothetical bias is unlikely to be a problem in our population. If instead the coefficient on the hypothetical dummy variable is positive and statistically significant, we conclude that positive hypothetical bias is likely to exist, and its average level is measured by the value of the coefficient.

[INSERT TABLE 1 HERE]

[11] Since households in our sample are spread across 40 western U.S. cities, we adjust CRP fees and income levels for differences in costs of living using the city comparison calculator at http://list.realestate.yahoo.com/re/neighbor/.

[12] More precisely, our pooled dataset is created by first including all CRP-V households that have effectively revealed their preferences via the decision of whether to participate in their community's voluntary CRP. The voluntary CRPs vary in terms of their monthly cost-of-living-adjusted fees (which are roughly between $1 and $5 per month) and whether they require sorting of the recyclables. Second, we include all CRP-H households (that valued hypothetical sorting and non-sorting CRPs) with random opening bids that were within the $1 to $5 interval. Although our WTP questions were set in a double-bounded, dichotomous-choice format (discussed in further detail below), we consider solely the opening bids for this exercise so that the design of the hypothetical decision mimics actual decisions as closely as possible. The pooled dataset contains 1,782 households – 994 CRP-V households (i.e., revealed-preference households) and 788 CRP-H households (i.e., stated-preference households).

[13] See List et al. (2004) and Carson et al. (2002) for two recent studies that measure how the degree of "social isolation" and "consequentialism," respectively, influence respondents' "real" responses in laboratory experiments.

The estimation results are shown in Table 1. The signs and significance levels of the demographic, program- and community-specific variables are similar to the results in Aadland and Caplan. We therefore focus on the result germane to this section; the sign and magnitude of the estimated coefficient on the dummy variable for whether the participation decision was hypothetical (CRP-H HOUSEHOLD). The coefficient is statistically significant and indicates that, all else equal, CRP-H households stated a WTP that is, on average, $2.76 more than CRP-V households. Stated in terms of likelihood of participation, the CRP-H households are 7.7 percent more likely to participate than CRP-V households. This suggests that positive hypothetical bias is present in our data.

4.2 Unconditional Cheap-Talk Effects

Does cheap talk mitigate the positive hypothetical bias in our data? In Table 2 we report average rates of hypothetical participation across opening bid levels, cheap-talk scripts, and types of CRP (actual or hypothetical). It is important to note that unlike the analysis in the previous section we are not interested in the actual participation decisions of CRP-V households. Instead, we focus on how cheap talk influences the stated participation decisions of both CRP-H and CRP-A households.

[INSERT TABLE 2 HERE]

Before discussing the effects of cheap talk, notice that, as expected, hypothetical participation rates (and, by implication, WTP) generally decline as the bid levels increase. More importantly, however, Table 2 indicates that respondents who were read either short or long cheap-talk scripts often stated they would participate at statistically different rates than those who did not receive cheap talk. The surprising result is not that cheap talk affects hypothetical participation decisions for curbside recycling, but rather that it is typically associated with *higher* rates of hypothetical participation. Furthermore, the strongest positive effect is for CRP-A households and households receiving the *longer* cheap-talk script with a reminder about substitutes and budget constraints. Thus, from the results in Table 2, it is clear that there is absolutely no evidence that cheap talk (either with or without a reminder about substitutes and budget constraints) is effective in mitigating the positive hypothetical bias known to exist in our data. To the contrary, cheap talk appears to exacerbate the bias!

We now offer a few potential explanations for this counterintuitive result. To begin, it could be argued that by including the word "landfills" in our long-script version we unwittingly created an environmental "flashpoint," triggering images of overflowing garbage dumps in the minds of respondents who then reacted by inflating their WTP responses. The problem with this hypothesis, however, is that it cannot completely explain our results. CRP-A households that receive short-scripted cheap talk (without the landfill reference) also tend to be more likely to participate (see Table 2) and, as we will see in Section 4.3, are more likely to state conditionally higher values than those not receiving any cheap talk.

A second possible explanation, provided by Cummings et al., is that cheap talk "might assure those doubting the hypothetical nature of the experiment that it is indeed hypothetical" and as a result encourages positive hypothetical bias. The difficulty with this explanation is in understanding why additional positive bias is elicited from some scripts but not others. A third possibility is that when they hear a neutral cheap-talk script stating that respondents' WTP responses "*are sometimes different*" as opposed to "*are sometimes higher*" than their actual WTP, survey respondents may be induced to rely more heavily on their own heuristics and inferences in an attempt to guess what type of bias the researchers have in mind. If this hypothesis is true, our results indicate that respondents may be guessing that we would expect them to correct this bias by stating higher WTP values.[14]

Each of these hypotheses and our empirical results suggest that we simply do not understand how the human cognitive process receives and then reacts to signals such as cheap talk. Several theories of how human cognition reacts to signals are discussed in Fischhoff (2002). As Fischhoff (in press) points out, artifacts (such as unexpected responses to cheap talk) could emanate from "the subtle ways that interviewers communicate their expectations." He (in press) also notes that "elicitation is a reactive measurement procedure...The process assumes that people sometimes need help, in order to understand what they believe and want. That help may include presenting a balanced selection of opinions, lest clients miss a critical perspective just because it did not occur to them at the time." It is possible that in erring on the side of conciseness, our short but balanced cheap-talk script provided insufficient detail regarding "selection of opinions," resulting in unpredictable effects on WTP.

As a final note, we acknowledge that our counterintuitive cheap-talk results could simply be due to systematic differences in the treatment and control groups. This seems unlikely, however, as the cheap-talk scripts were assigned randomly across such a large number of households and communities. Nevertheless, for the sake of thoroughness, we now turn to a conditional analysis of cheap talk's effects.

4.3 Conditional Cheap-Talk Effects

Our conditional estimates are based on the double-bounded, dichotomous-choice (DBDC) model first introduced by Carson et al. (1986). We use maximum likelihood to estimate a model

[14] Alternatively, respondents may be reacting to what they believe the university-sponsored interviewer wants to hear, thus creating a "social-desirability bias" in favor of curbside recycling. We thank John Loomis for this insight.

that incorporates the responses to both the opening and follow-up bids. As in the participation probit model in Section 4.1, we control for a number of demographic and community attributes.[15] Based on our earlier work and List, we also estimate models that allow for differential cheap-talk effects across household characteristics.

WTP questions set in the DBDC format elicit a household's WTP through a sequence of yes-or-no valuation questions. The first question is: "Would you be willing to pay $\$\tau$ for the service?" The opening bid τ is chosen randomly from the integers \$2 through \$10, based on the range of actual household fees in our sub-sample of communities with an existing CRP. Based on her response to the opening bid, the respondent is then asked a similar follow-up question, but with a larger bid, $\tau_H = 2\tau$, if she answered "yes" (i.e., willing to pay at least τ for the service) or a smaller bid $\tau_L = 0.5\tau$ if she answered "no" (i.e., unwilling to pay τ for the service). Those who answer "no" to the first two questions are then asked whether they would participate in a program that is free of charge. Based on the responses to the opening bid and follow-up questions, the respondent's latent WTP may therefore be placed in one of five regions: $(-\infty,0)$, $[0,\tau_L)$, $[\tau_L,\tau)$, $[\tau,\tau_H)$ or $[\tau_H, \infty)$.

There is a growing literature concerned with incentive incompatibility issues arising from use of the DBDC format (c.f., Alberini et al. 1997, Carson et al. 2000, Cameron et al. 2002, and DeShazo 2002). Recently, Aadland and Caplan (2004) have proposed a modified version of Whitehead's (2002) random-effects probit model to test and control for both starting-point bias and incentive incompatibility in iterative-question formats. When applied to our data, the modified random-effects model indicates the existence of starting-point bias and a small amount of incentive incompatibility. However, the mean WTP estimates for the two models (one

[15] The estimated coefficients on these control variables are generally similar in sign and significance to the results in Table 1, and therefore are omitted.

controlling for starting-point bias and incentive incompatibility and one not) are very similar. As a result, we report the results from the latter model.[16]

We posit that a household's true WTP (WTP*) is represented by the equation

$$WTP_i^* = X_i\beta + \varepsilon_i, \tag{1}$$

where X_i is a row vector of household-, program-, and community-specific control variables, β is a corresponding column vector of coefficients, and ε_i is a normally distributed error term for households i = 1,…,n. We also allow for possible heteroscedasticity by modeling the variance of the WTP error term as

$$\sigma_i^2(Z_i\gamma) = \exp(Z_i\gamma), \tag{2}$$

where Z_i is a row vector of variables related to the disturbance variances and γ is a column vector of parameters.

The results from this exercise are reported in Table 3. The first row of Table 3 shows the effect of short and long cheap talk across all CPR-H and CRP-A households. The remaining rows report the results of the two cheap-talk scripts interacted with certain demographic groups. Consistent with the unconditional cheap-talk results in the previous section, all of the coefficients are either positive or not statistically different from zero. Overall, survey respondents are more sensitive to cheap talk when they are valuing an actual CRP and when cheap talk is accompanied by a reminder about substitutes and budget constraints.[17]

[INSERT TABLE 3 HERE]

Furthermore, CRP-A households with at least one member holding a Ph.D. or equivalent professional degree and who received short (long)-scripted cheap talk are, all else equal, willing to pay approximately $1.07 (1.38) more per month for an actual CRP than their counterparts who

[16] The results from the former model are available from the authors upon request.
[17] Consistent with Bulte et al., we also find no evidence that cheap talk is related to the variance of the errors.

did not receive short (long)-scripted cheap talk. Similarly, respondents who belong to an environmental organization and who received the short script are willing to pay an additional $1.54 per month, on average, while those who received the long script are only willing to pay an additional $1.45 per month. In sum, our results indicate that cheap-talk statements may be ineffective or even counter-productive in mitigating hypothetical bias, with the degree of the problem varying across types of households.

5. Conclusion

The evidence from our CVM survey draws into question the efficacy of cheap talk as a reliable and robust *ex ante* correction mechanism for positive hypothetical bias. Although initial research has shown that a long-scripted version of cheap talk can be effective in eliminating this bias in lab experiments and field auctions, shorter and more neutrally worded scripts appropriately tailored for phone interviews have clearly demonstrated a high degree of sensitivity to script length and content. Indeed, we first establish the existence of positive hypothetical bias in our data by contrasting revealed- and stated-preference information, but then find that our neutral cheap-talk scripts actually exacerbate the problem. Moreover, the degree of exacerbation seems to increase with script length and with respect to household characteristics (e.g., education, environmental advocacy, etc.) that are likely to be systematically related to the degree of observed hypothetical bias. Because WTP responses are so sensitive to script length and content, CVM practitioners should use caution in relying on cheap-talk statements to mitigate hypothetical bias.

A potential criticism of our cheap-talk design is that we did not *a priori* establish a baseline degree of hypothetical bias in our data. Had we known the extent of positive hypothetical bias in our data beforehand, we could have chosen a script informing respondents of the direction and (possibly the magnitude) of the bias, rather than using a neutral script. CT, List, and Aadland and Caplan (2003) report some success with these types of directed scripts. However, the primary attraction of cheap talk as an *ex ante* control for hypothetical bias is its apparent generalizability (in the form of a standard script) to CVM studies across a wide array of non-market goods and services. Our findings suggest that standardized cheap-talk scripts can produce undesirable results. As a result, we feel caution is warranted in using cheap talk to correct for hypothetical bias until we better understand how the length and content of cheap talk statements influence the cognitive processes of survey respondents.

Appendix

Explanatory Variables	Description
Ethical Duty	Do you feel an ethical duty to recycle to help the environment? 1 = yes, 0 = no.
Monetary	Are you motivated to recycle in order to save money? For example, are you able to use a smaller garbage container because you recycle or you get money for your aluminum cans? 1 = yes, 0 = no.
Primarily Ethics	Which one would most encourage your household to recycle: an ethical duty to help the environment, or saving money? 1 = ethical duty, 0 = save money.
Dropoff Distance	Distance in miles to the nearest dropoff site.
Dropoff User	For households that have used dropoff facilities in the last 12 months, how often do you take recyclable materials to the dropoff center? 1 = always or often, 0 = sometimes or rarely.
Young	1 if 18<Age<35, 0 otherwise.
Old	1 if 65<Age, 0 otherwise.
Male	1 = male, 0 = female.
High School	What is the highest level of education attained by any member of your household? 1 = high school graduate, 0 = otherwise
Associates	1 = associates degree, 0 = otherwise
Bachelors	1 = bachelors degree, 0 = otherwise
Masters	1 = masters degree, 0 = otherwise
Ph.D.	1 = Ph.D. or equivalent professional degree, 0 = otherwise
Household Size	Number of adults currently living in the household who are older than 18 years of age, other than the respondent.
Environmental Organization	Does anyone in your household belong to an environmental club, group, or organization? 1 = yes, 0 = no.
Med Income	1 if $35K/yr<Household Income<$75K/yr, 0 otherwise
High Income	1 if Household Income>$75K/yr, 0 otherwise
Employed	Is the adult with the highest income currently working for pay, either full time or part time? 1 = yes, 0 = no.
Retired	Is the adult with the highest income currently retired? 1 = yes, 0 = no.
Short Cheap Talk	1 = received short cheap-talk statement, 0 otherwise.
Longer Cheap Talk	1 = received longer cheap-talk statement, 0 otherwise.
Sorting Required	1 = CRP requires some sorting of recyclable materials, 0 otherwise.
Polite	1 if polite initial refusal, 0 otherwise.
Angry	1 if angry initial refusal, 0 otherwise.
Certainty	1 = certainty of response to the last WTP question ≥ 90%, 0 otherwise
Landfill Visit	Has anyone in your household ever visited your community's landfill? 1 = yes, 0 = no.
Landfill Distance	Distance to nearest landfill in miles.
Landfill Distance Spline	Max{Landfill Distance – 2, 0}.
CRP-H Household	1 = hypothetical CRP household, 0 = otherwise.

References

Aadland, D., Caplan, A.J., 2004. Incentive incompatibility and starting-point bias in iterative valuation questions: comment. Land Economics 80(2), 312-315.

Aadland, D., Caplan, A.J., 2003. Willingness to pay for curbside recycling with detection and mitigation of hypothetical bias. American Journal of Agricultural Economics 85(2), 492-502.

Adamowicz, W., Louviere, J., Williams, M., 1994. Combining revealed and stated preference methods for valuing environmental amenities. Journal of Environmental Economics and Management 26, 271-292.

Alberini, A., Kanninen, B., Carson, R.T., 1997. Modeling response incentive effects in dichotomous choice contingent valuation data. Land Economics 73, 309-324.

Andreoni, J., 1990. Impure altruism and donations to public goods: a theory of warm-glow giving. The Economic Journal 100, 464-477.

Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R., Schuman, H., 1993. Report of the NOAA panel on contingent valuation. Federal Register 58(10), 4601-14.

Brown, T. C., Ajzen, I., Hrubes, D., 2003. Further tests of entreaties to avoid hypothetical bias in referendum contingent valuation. Journal of Environmental and Economics Management 46, 53-361.

Bulte, E., Gerking, S., List, J.A., de Zeeuw, A., 2003. "The effect of varying the causes of environmental problems on stated values: evidence from a field study." Unpublished manuscript, University of Central Florida, Orlando.

Cameron, T.A., Poe, G.L., Ethier, R.G., Schulze, W.D., 2002. Alternative non-market value-elicitation methods: are the underlying preferences the same? Journal of Environmental Economics and Management 44, 391-425.

Carson, R.T., Hanneman, W., Mitchell, R.C., 1986. Determining the demand for public goods by simulating referendums at different tax prices." Unpublished manuscript, University of California, San Diego.

Carson, R.T., Flores, N.E., Martin, K.M., Wright, J.L., 1996. Contingent valuation and revealed preference methodologies: comparing the estimates for quasi-public goods. Land Economics 72, 80-89.

Carson, R.T., Groves, T., Machina M., 2000. Incentive and informational properties of preference questions. Unpublished manuscript, University of California at San Diego.

Carson, R.T., Groves, T., List, J.A., Machina, M., 2002. Probabilistic influence and supplemental benefits: a field test of the two key assumptions underlying stated preferences." Draft paper for presentation at the World Congress of Environmental and Resource Economists, June.

Cummings, R.G., Taylor L.O., 1999. Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. American Economic Review 89(3), 649-666.

Cummings, R.G., Harrison, G.W., Osborne, L.L., 1995. Can the bias of contingent valuation be reduced? evidence from the laboratory. Economics Working Paper B-95-03, Division of Research, College of Business Administration, University of South Carolina.

DeShazo, J.R., 2002. Designing transactions without framing effects in iterative question formats. Journal of Environmental Economics and Management 43(3), 360-385.

Diamond P.A., Hausman, J.A., 1994. Contingent valuation: is some number better than none?" Journal of Economic Perspectives 8(4), 45-64.

Dickie, M., Fisher, A., Gerking, S., 1987. Market transactions and hypothetical demand data: a comparative study. Journal of the American Statistical Association 82(397), 69-75.

Fischhoff, B., 2002. Cognitive processes in stated preference methods. Forthcoming in: Mäler, K.-G., Vincent J. (Eds.). Handbook of Environmental Economics, Vol. II. Amsterdam: North-Holland.

Freund, J.E., 1992. Mathematical Statistics, 5th Edition. Englewood Cliffs: Prentice Hall.

Haab, T., Huang, J.C., Whitehead, J., 1999. Are hypothetical referenda incentive compatible: a comment. Journal of Political Economy 107(1), 186-196.

Hanneman, M., 1994. Valuing the environment through contingent valuation. Journal of Economic Perspectives 8(4), 19-43.

Harrison, G.W., 2002. Experimental economics and contingent valuation. Unpublished manuscript, Department of Economics, Moore School of Business, University of South Carolina.

Harrison, G.W., Beekman, R.L., Brown, L.B., Clements, L.A., McDaniel, T.M., Odom, S.L., Williams, M., 1999. Environmental damage Assessment with hypothetical surveys: the calibration approach. In: Boman, M., Brännlund, R., Kriström, B. (Eds.). Topics in Environmental Economics. Amsterdam: Kluwer Academic Press, 217-240.

Kahneman, D., Tversky, A., 1984. Choices, values, and frames. American Psychologist 39, 341-350.

List, J.A., 2001. Do explicit warnings eliminate the hypothetical bias in elicitation procedures? evidence from field auction experiments. American Economic Review 91(5), 1498-1507.

List, J.A., Shogren, J.F., 1998. Calibration of the difference between actual and hypothetical evaluations in a field experiment. Journal of Economic Behavior and Organization 37(2), 193-205.

List, J.A., Berrens, R.P., Bohara, A.K., Kerkvliet, J., 2004. Examining the role of social isolation on stated preferences. Forthcoming in the American Economic Review.

Loomis, J.B., Caban, A.-G., Gregory, R., 1994. Substitutes and budget constraints in contingent valuation. Land Economics 70(4), 499-506.

Loomis, J.B., Brown, T., Lucero, B., Peterson, G., 1996. Improving validity experiments of contingent valuation methods: results of efforts to reduce the disparity of hypothetical and actual willingness to pay. Land Economics 72(4), 450-461.

Murphy, J.J., Stevens, T., Weatherhead, D., 2003. An empirical study of hypothetical bias in voluntary contribution contingent valuation: does cheap talk matter? Unpublished manuscript, University of Massachusetts, Amherst.

Neil, H.R., 1995. The context for substitutes in CVM studies: some empirical observations. Journal of Environmental Economics and Management 29(3), 393-397.

Nestor, D.V., 1998. Policy evaluation with combined actual and contingent response data. American Journal of Agricultural Economics 80, 264-276.

Poe, G.L., Clark, J., Schulze, W., 2002. Provision point mechanisms and field validity tests of contingent valuation. Environmental and Resource Economics 23, 105-131.

Whitehead, J.C., 2002. Incentive incompatibility and starting-point bias in iterative valuation questions. Land Economics 78(2), 285-297.

Table 1. Single-Bounded Probit Model for CRP Participation

Explanatory Variables	Estimates		Summary Statistics		
	Coefficient	P –Value	N	Mean	Std. Dev.
Ethical Duty	4.451***	0.000	1759	0.869	0.340
Monetary	-2.579***	0.000	1757	0.484	0.500
Primarily Ethics	1.564***	0.005	753	0.578	0.494
Dropoff Distance	0.082*	0.096	1248	3.843	3.557
Dropoff User	-0.630	0.120	842	0.618	0.237
Young	1.044**	0.011	1715	0.283	0.451
Old	-1.386**	0.034	1715	0.118	0.323
Male	-0.144	0.343	1768	0.386	0.487
High School	1.322	0.130	1755	0.133	0.338
Some College	1.184	0.153	1755	0.165	0.372
Associates	1.784*	0.075	1755	0.102	0.303
Bachelors	2.193**	0.032	1755	0.308	0.462
Masters	2.672**	0.017	1755	0.189	0.391
Ph.D.	2.508**	0.030	1755	0.079	0.270
Household Size	0.016	0.466	1761	1.109	0.922
Environmental Organization	1.747***	0.007	1744	0.094	0.292
Med Income	-0.026	0.480	1563	0.406	0.491
High Income	0.245	0.330	1563	0.364	0.481
Employed	0.827	0.140	1748	0.815	0.389
Retired	1.452*	0.068	1733	0.124	0.329
Short Cheap Talk	0.641	0.159	1768	0.153	0.360
Longer Cheap Talk	1.941***	0.004	1768	0.145	0.352
Sorting Required	-1.629***	0.001	1768	0.420	0.494
Polite	-1.331***	0.008	1768	0.137	0.344
Landfill Visit	0.159	0.352	1740	0.556	0.497
Landfill Distance	-4.067***	0.007	1238	10.562	7.798
Landfill Distance Spline	3.946***	0.008	1238	8.594	7.761
Certainty	-0.3722	0.180	1746	0.753	0.432
CRP-H Household	2.756***	0.000	1768	0.442	0.497
Heteroscedasticity Variables					
Constant	2.906***	0.000	1768	1.000	0.000
(Opening) Bid	0.138	0.111	1768	2.847	1.028

Notes: ***, **, and * refer to statistical significance at the 1, 5 and 10 percent levels respectively. The dependent variable is participation in a CRP. The estimates for the constant term, as well as binary variables for "don't know" and missing responses are not shown. Overall sample size = 1768. The varying number of observations (N) under the descriptive statistics reflects "don't know" and missing responses. The likelihood ratio statistic for the null hypothesis that all the coefficients on the explanatory variables are jointly equal to zero is LR = 336.61 and is statistically significant at the 1% level.

Table 2. Unconditional Participation Rates across Cheap-Talk Scripts

		CRP-H Households				CRP-A Households			
		No C-Talk	Short C-Talk	Long C-Talk	N	No C-Talk	Short C-Talk	Long C-Talk	N
COLA-adjusted Opening Bids ($/month)	≤ 2	58.3	63.0	63.0	109	61.9	75.7**	82.3***	195
	(2, 3]	64.0	64.9	77.5**	243	65.9	71.2	65.8	306
	(3, 4]	49.3	72.6***	77.9***	234	56.9	66.0	70.6**	298
	(4, 5]	61.8	55.7	65.4	233	39.8	59.6***	60.3***	328
	(5, 6]	60.0	52.2	62.4	222	44.9	51.6	55.6*	319
	(6, 7]	54.1	44.9	54.8	192	33.0	39.8	43.4*	272
	(7, 8]	39.1	38.2	36.8	213	24.4	32.4	41.7**	221
	(8, 9]	44.4	47.6	41.7	168	33.3	28.6	35.6	196
	(9, 10]	47.2	31.7	43.1	164	23.3	32.6	41.9**	132
	> 10	30.8	43.2	41.9	132	19.2	37.1*	30.0	91
Totals					1910				2358

Notes: *, ** and *** indicate significantly different than No C-Talk at the 10, 5 and 1 percent significance levels respectively. Statistical tests for the differences in proportions are calculated as described in Freund (1992, p. 414-6). CRP ≡ Curbside Recycling Program. C-Talk ≡ Cheap Talk. COLA ≡ Cost of Living Adjusted.

Table 3. Conditional Estimates of Cheap-Talk Scripts on Maximum WTP

Interactive Terms	Cheap-Talk Coefficients			
	CRP-H Households		CRP-A Households	
	Short C-Talk	Long C-Talk	Short C-Talk	Long C-Talk
None	0.188 (0.256)	0.584** (0.259)	0.486** (0.227)	0.794*** (0.223)
Young	0.403 (0.437)	0.647* (0.436)	0.796** (0.433)	1.037*** (0.419)
Female	-0.286 (0.342)	0.272 (0.337)	0.582** (0.290)	0.973*** (0.287)
Env. Org.	0.674 (0.989)	0.789 (1.018)	1.539** (0.754)	1.454** (0.758)
Bachelors	0.048 (0.459)	0.751* (0.482)	1.051* (0.748)	-0.100 (0.748)
Ph.D.	0.485 (0.944)	0.578 (0.952)	1.074* (0.799)	1.383** (0.767)
Sorting Required	0.237 (0.366)	0.720** (0.372)	0.323 (0.363)	0.819** (0.363)
Ethical Duty	0.312 (0.285)	0.579** (0.284)	0.435** (0.238)	0.825*** (0.235)
Certainty ≥ 90	0.307 (0.316)	0.712** (0.322)	0.631** (0.277)	0.935*** (0.270)

Notes: *, ** and *** indicate significantly different than No C-Talk at the 10, 5 and 1 percent significance levels respectively. Standard errors are in parentheses. CRP ≡ Curbside Recycling Program. WTP ≡ Willingness To Pay. C-Talk ≡ Cheap Talk. Sample Size = 4253.