



College of Business
and Public Policy
UNIVERSITY of ALASKA ANCHORAGE

Department of Economics Working Paper
WP 2010-06
March 2010

A Comparison of Induced Value and Home-Grown
Value Experiments to Test for Hypothetical Bias
in Contingent Valuation

James J. Murphy
University of Alaska Anchorage

Thomas H. Stevens
University of Massachusetts Amherst

Lava Yadav
University of Massachusetts Amherst

UAA DEPARTMENT OF ECONOMICS
3211 Providence Drive
Rasmuson Hall 302
Anchorage, AK 99508

<http://econpapers.uaa.alaska.edu>

Revised March 2010

A Comparison of Induced Value and Home-Grown Value Experiments to Test for Hypothetical Bias in Contingent Valuation

JAMES J. MURPHY ^{a,b,*}

THOMAS H. STEVENS ^b

LAVA YADAV ^b

^a Department of Economics, Rasmuson Hall, 3211 Providence Drive, University of Alaska Anchorage, Anchorage, AK 99508

^b Department of Resource Economics, Stockbridge Hall, 80 Campus Center Way, University of Massachusetts-Amherst, Amherst, MA, 01003, USA

* Corresponding author.
Phone: 907-786-1936
Fax: 907-786-4115
Email: murphy@uaa.alaska.edu

This paper has not been submitted elsewhere in identical or similar form, nor will it be during the first three months after its submission to the Publisher.

A Comparison of Induced Value and Home-Grown Value Experiments to Test for Hypothetical Bias in Contingent Valuation

Abstract: This study tests the hypothesis that hypothetical bias may not be related to value elicitation; rather it may be a value formation problem. When participants are asked to indicate their willingness to pay for an induced value good, we find no evidence of hypothetical bias for three different commodity types (public good, private good, and publicly provided private good). However, when these same subjects are asked to value homegrown goods with no pre-assigned induced value using the same elicitation mechanism, hypothetical values are roughly double actual payments in all three cases. These results support the hypothesis that the process of forming values in a homegrown setting may be a key contributor to hypothetical bias.

Keywords: contingent valuation, hypothetical bias, experiments, induced values, home-grown values.

JEL Classification: C91, H41, Q26, Q28.

1 **1. Introduction**

2 Contingent valuation (CV) is one of the primary tools for estimating the value of non-market
3 goods. The methodology can play a key role in policy decisions that rely on benefit-cost
4 analysis, and as the *Exxon Valdez* case highlighted, CV can be instrumental in determining
5 penalties for environmental damages. Despite CV's prominent role, its use as a decision making
6 tool is often quite controversial. Because CV surveys are hypothetical in both the payment for
7 and provision of the good, it is impossible to determine whether the choices made in a
8 hypothetical survey are consistent with how the individual would behave if actually given an
9 opportunity to do so. This inconsequential nature of CV surveys, i.e., the lack of a salient
10 economic commitment, can lead to hypothetical bias in which CV overestimates the true
11 economic value of the good.

12 Although the existence of hypothetical bias is well-documented (List and Gallet 2001,
13 Murphy, *et al.* 2005a, Harrison and Rutström 2008), its underlying causes are not well
14 understood (Murphy and Stevens 2004). Without a theory to help explain why hypothetical bias
15 persists, the general applicability of calibration techniques designed to offset these effects may
16 be limited. The cheap talk approach (Cummings and Taylor 1999), for example, is a popular
17 calibration technique in which the hypothetical bias problem is described to subjects. Yet, the
18 effectiveness of this approach may be sensitive to key variables such as script length (Poe, *et al.*
19 2002; Aadland and Caplan 2003), subject experience (List 2001; Lusk 2003; Aadland and
20 Caplan 2003) and payment amounts (Brown, *et al.* 2003; Murphy, *et al.* 2005b). In fact, Aadland
21 and Caplan 2006 find that a neutral cheap talk script can actually exacerbate hypothetical bias.¹

¹ See Murphy and Stevens 2004, and Horowitz *et al.* forthcoming, for a more comprehensive discussion of this literature.

1 The vast majority of contingent valuation hypothetical bias experiments elicit subjective,
2 homegrown values, rather than experimenter-controlled, pre-assigned induced values. Unlike
3 induced values, the researcher cannot know these homegrown values with certainty. Because of
4 this, when researchers observe that values in a hypothetical payment scenario are higher than the
5 corresponding treatment with actual, consequential payments, it is impossible to know for sure
6 whether hypothetical values are overstated or actual values are understated (or possibly a
7 combination of the two). Researchers typically make the reasonable assumption that the
8 responses in the real settings accurately represent the true economic value, yet it is entirely
9 possible that the reverse is true. For example, when payments for public good provision are
10 consequential, responses could be biased downward due to factors like free-riding or the desire
11 to only pay one's fair share. In the case of private goods, both Harrison, *et al.* 2004 and Murphy
12 and Stevens 2004 hypothesize that responses to actual payment questions for private goods may
13 be censored by the market price. The linkage between value elicitation for public and private
14 goods is particularly important considering that many experimental studies of non-market
15 valuation techniques use private goods, and as with public goods, hypothetical bias is often
16 observed.

17 Much is known about the presence of hypothetical bias with homegrown goods, even
18 though the underlying causes are not fully understood. There is a recent increase in the use of
19 induced value experiments to shed light on the hypothetical bias observed when eliciting
20 homegrown values. In induced value experiments, the true underlying value is known and
21 therefore deviations from this value are more readily identified. In an induced value experiment
22 that uses a referendum for provision of a public good, Taylor, *et al.* 2001 find no evidence of
23 hypothetical bias in aggregate, even though they observe errors at the individual level. They

1 conclude that hypothetical bias is not a value elicitation problem, but rather that value formation
2 may be at the heart of hypothetical bias. Their conclusion is certainly reasonable, and dovetails
3 nicely with the literature about value uncertainty.² However, respondents misvoted—that is,
4 voted in a manner inconsistent with their induced values—over 16% of the time in both the real
5 and hypothetical payment treatments. It turns out that even though roughly one in six responses
6 was a “decision error,” there was no hypothetical bias in aggregate.

7 Even in the absence of any hypothetical bias, it is still possible that the WTP estimates
8 could be inaccurate. This is precisely what Mitani and Flores (2009) observe. They find no
9 evidence of hypothetical bias in aggregate for a threshold public good with an open-ended
10 elicitation procedure. However, only 8% of respondents truthfully revealed their true value, 75%
11 underreported their values and 17% overreported, leading to WTP estimates that are significantly
12 below induced values. Vossler and McKee 2006 investigate value uncertainty as a possible
13 cause of hypothetical bias by testing four elicitation mechanisms: dichotomous choice,
14 dichotomous choice with a follow-up certainty question, payment card, and multiple-bounded
15 discrete choice. They find no evidence of hypothetical bias in aggregate with any of these
16 mechanisms, although similar to Taylor *et al.*, they do find evidence of decision errors at the
17 individual level. There are also two studies that use induced values to test the properties of
18 contingent valuation when individual costs are uncertain (Burton, *et al.* 2003; Polomé 2003), but
19 neither has an inconsequential payment condition to draw inferences about hypothetical bias.

² The relevant literature regarding uncertain preferences includes Opaluch and Segerson 1989, Gregory, *et al.* 1995, Li and Mattsson 1995, Gregory and Slovic 1997, Wang 1997. Subsequent studies which use a respondent’s self-reported level of uncertainty as an *ex post* calibration mechanism include Champ, *et al.* 1997, Blumenschein, *et al.* 1998, Johannesson, *et al.* 1998, Ethier, *et al.* 2000, Champ and Bishop 2001, Poe, *et al.* 2002. For a review of this literature, see Murphy and Stevens 2004.

1 Interestingly, the three aforementioned induced value studies report no evidence of
2 hypothetical bias in aggregate for a public good provided via either a referendum or a voluntary
3 contribution with a provision point. However, Cherry *et al.* 2004 do report the presence of
4 hypothetical bias for a private good in an induced value experiment that uses a theoretically
5 incentive-compatible second price auction to elicit values (we revisit this result later). Moreover,
6 there was substantial individual decision error in their no outside option treatment (which is the
7 closest parallel to our study).

8 This small but growing literature suggesting the absence of hypothetical bias in aggregate
9 in induced value settings, but a nontrivial level of biased responses persists at the individual
10 level, even with theoretically incentive compatible elicitation mechanisms. In most induced
11 value studies, the individual decision errors have been consistent between real and hypothetical
12 payments, leading to no hypothetical bias in aggregate. Nevertheless, these individual errors
13 could still lead to inaccurate WTP estimates and also raise concerns about the conditions under
14 which these individual errors arise.

15 Both Taylor *et al.* and Vossler and McKee focused solely on eliciting induced values for
16 public goods. Our experiments extend this line of research by linking not only the public and
17 private good literatures, but also the induced value and homegrown value literatures, so we can
18 better understand how inferences from one context can provide insights into another. We do not
19 observe hypothetical bias in aggregate in induced value experiments for any good type, but when
20 subjects are asked to value homegrown goods, hypothetical bias arises. This provides some
21 support for Taylor *et al.*'s 2001 conclusion that hypothetical bias is a value formation problem,
22 not one of value elicitation. Like Taylor *et al.*, we also observe the presence of individual
23 decision errors, or misvotes, in induced value referenda. However, when subjects are given a

1 purely private decision with no strategic interaction or potential for other regarding behavior,
2 these individual decision errors disappear. This would suggest that value elicitation, particularly
3 in a referendum, may affect responses at the individual level.

4

5 **3. Experimental Design**

6 A total of 169 students were recruited from the general student population at the University of
7 Massachusetts, Amherst. Subjects were informed about the experiments through flyers that were
8 placed around campus and a recruiting table that was set up in the campus center. Subjects were
9 paid \$5 for showing up on time and were told they would be given an opportunity to earn
10 additional money during an experiment that would last no more than 90-minutes (actual
11 experiment time was less than one hour).

12 Our 3×2 (type of good × payment type) experimental design is shown in Table 1. Each of
13 the six cells represents a unique treatment. Each subject participated in only one of the six
14 treatments. The data were collected on six separate days, one treatment each day. The three rows
15 indicate the three good types (private, publicly-provided private, or public). As shown in Table
16 1, the switch from a private to public good context entails changes to both the good used to elicit
17 homegrown values and the elicitation mechanism. To control for potential confounding effects,
18 we also included a publicly-provided private good as an intermediate step between the two good
19 types. The payment type columns reflect the two payment conditions (real or hypothetical). The
20 real and hypothetical treatments differed only in that outcomes in the former were consequential,
21 whereas in the latter subjects received a pre-announced fixed payment regardless of the
22 experiment results. The instructions in the hypothetical payment treatments were modified
23 slightly to reflect the inconsequential nature of the payments. Subjects in these hypothetical

1 treatments were told: “Each of you will receive exactly \$13.50 at the end of the experiment today
2 (in addition to the \$5 you already received for participating), regardless of your decisions and the
3 outcomes... Although your earnings in today’s experiment are fixed at \$13.50, we ask you to
4 suppose that your earnings were based on your decisions and the auction outcomes...” The
5 \$13.50 fixed earnings in the hypothetical payment treatments equal the expected earnings in the
6 real payment treatments.³

7 In each of the six treatments in Table 1, subjects participated in a total of five rounds. In
8 each round of a particular treatment, the type of good, payment type and elicitation mechanism
9 remained constant; the only factor that varied across the five rounds for a particular subject was
10 whether induced or homegrown values were elicited. As shown in Table 2, the first two rounds
11 were practice rounds using induced values to familiarize subjects with the instructions and the
12 elicitation mechanism.⁴ After both practice rounds were completed, the transaction prices for
13 each practice round were determined and results announced. The results from the practice
14 rounds did not count towards an individual’s earnings and are not included in the data. After
15 completion of the two practice rounds (P1 and P2), stage 1 (rounds 1 and 2) was introduced.
16 Like the practice rounds, stage 1 also used induced values. Stage 2 (round 3) followed the same
17 procedures as the previous rounds with the exception that induced values were no longer used.
18 Instead, homegrown values for a specific good (coffee mug or a contribution to Heifer
19 International) were elicited.

³ It turns out that the additional earnings in the real payment treatments were higher, averaging \$17.43 ($\sigma=2.78$). This difference between actual and expected earnings is due to the transaction prices that were randomly selected using a BDM-style elicitation mechanism described shortly.

⁴ Induced values were described as the purchase of a token in the private and publicly provided private good experiments, and a benefit from a project in the public good experiments.

1 Hence, for each subject, we have three WTP observations (two induced values from stage
2 one, and one homegrown value from stage two), all of which were elicited in same manner under
3 the same payment condition (i.e., real or hypothetical). This design allows us to make within-
4 subject comparisons of outcomes in induced vs. homegrown values settings while holding the
5 payment condition and good type constant, and between-subject tests for hypothetical bias
6 holding constant the type of good and value type (i.e., induced or homegrown).

7 ***Private good.*** In the two private good treatments (one each for real and hypothetical
8 payment), values were elicited using a slightly modified version of the Becker-DeGroot-
9 Marschak (BDM) mechanism (Becker *et al.* 1964).⁵ For both the induced value (stage one) and
10 homegrown value (stage two) elicitation, subjects received a purchase offer slip that contained
11 each of the 15 possible prices, ranging between \$1 and \$15 in whole dollar increments. Table 3
12 shows an example of the purchase offer slip for a subject with an induced value of \$4.50 in the
13 private good real payment treatment. For each of these possible prices, subjects indicated
14 whether they were willing to make a purchase at that price. Bids had to be consistent in that if
15 they indicated “yes” to a particular price, then they had to indicate yes to all lower prices. With
16 this payment card approach, an individual’s maximum WTP lies between the highest value to
17 which she indicated “yes” and the next highest amount. The transaction price was determined by
18 randomly selecting a marble from a basket containing 15 marbles numbered 1 through 15; the
19 number on the marble was the price. If the participant indicated “yes” to making a purchase at

⁵ Horowitz (2006) shows that the BDM is not theoretically incentive compatible for some behavioral models which are outside an expected utility framework. The empirical results are mixed. Both Irwin *et al.* (1998) and Vossler and McKee (2006) provide experimental evidence supporting the demand revealing properties of the BDM, and Plott and Zeiler (2005) cite the BDM’s incentive compatibility, particular with subjects who are trained in a mechanism which might otherwise be unfamiliar. On the other hand, Noussair *et al.* (2004) find that the second-price auction is a more effective at eliciting willingness-to-pay than the BDM. Lusk and Rousu (2006) also found that the second-price and random *n*th price auctions are more accurate than the BDM on average.

1 the randomly selected price, then a transaction occurred at that price. Earnings were the
2 difference between the induced value of the token and the transaction price. If the participant
3 indicated “no” for the selected price, then no transaction occurred and earnings were zero. This
4 process was repeated for each participant.

5 The first stage used induced values and consisted of two separate individual decision
6 tasks using tokens with pre-assigned values. Before the stage one tasks, all subjects were told
7 they had \$10 in an initial cash balance and that any earnings would be added to this. All subjects
8 had an induced value of \$4.50 in the first round and \$11.50 in the second. These values were
9 private information and subjects were told to make no assumptions about the others’ token
10 values. Assuming truthful demand revelation, expected earnings in the first round are \$0.53 and
11 \$4.03 in the second.

12 The second stage (round 3) elicited homegrown values. Each subject was presented with
13 an opportunity to purchase one 16-ounce stainless steel travel coffee mug with the logo of a
14 popular on-campus student-run coffee shop. The mug was described as follows:

15 We will sell these 16-ounce stainless steel travel mugs from People’s Market, which I will now
16 pass around for you to see. As many of you probably know, People’s Market is a non-profit
17 student run co-op located in the Student Union. This mug entitles you to a 20-cent discount on
18 each of your future coffee purchases at People’s Market. This discount only applies to coffee
19 purchases; it does not apply to any other beverage, including tea. People’s Market does not have a
20 refund policy, so if you buy a mug you will not be able to return it. (If it is defective, you will be
21 able to exchange it for a replacement).⁶
22

23 Other than the change from induced values to homegrown values, the second stage proceeded in
24 exactly the same manner as the first. After the second stage was completed, all three prices (two
25 for stage one, and another for the second stage) were determined using the BDM mechanism.

⁶ The no-refund policy removes the possibility of purchasing the mug with the sole intent of returning it for the \$10 retail price.

1 Outcomes were announced at the end of the experiment to avoid the possibility that later
2 decisions could be conditioned on prior results.

3 *Public good.* In the two public good treatments, values were elicited through a BDM-
4 style referendum that was modified to parallel the mechanism used for private goods as closely
5 as possible (Vossler and McKee, 2006). Subjects received a voting slip similar to the purchase
6 offer slip in Table 3. For each of the 15 possible “project costs,” subjects had to indicate how
7 they would vote in a referendum that required all participants to pay this amount, and as with the
8 BDM, votes had to be consistent. The project cost was determined by randomly selecting a
9 marble from a basket containing 15 marbles numbered 1 through 15; the number on the marble
10 was the cost. After determining the cost, we tallied the votes for that cost. If more than half the
11 subjects voted yes to the randomly selected cost, then the project was implemented and each
12 subject paid that amount.

13 The first stage used induced values and was presented as a pair of group projects from
14 which each participant received some pre-assigned benefit. The initial cash and induced values
15 were the same as the private good BDM. In the second stage (with homegrown values), the
16 induced values were replaced with a referendum on whether the group would make a
17 contribution to Heifer International to purchase a flock of chickens for needy families. Subjects
18 were given promotional materials from Heifer International, and the referendum was described
19 as follows:

20 Heifer International is a nonprofit, humanitarian organization dedicated to ending world hunger
21 and saving the earth by providing livestock, trees, training and other resources to help poor
22 families around the globe become self-reliant. Since 1944, Heifer International has worked
23 directly with millions of families in more than 125 countries to alleviate hunger and increase self-
24 reliance. The livestock given to poor families (such as cows, llamas, goats, chickens, etc.), provide
25 milk, eggs, plowing power and other benefits that can mean improved nutrition, higher income,
26 education for children, health care, improved housing and a new way of life. Heifer requires all
27 livestock recipients to “Pass on the Gift” by sharing offspring of their livestock with others in
28 need. This common sense approach leads to sustainable development of one family at a time.

1 You will vote to decide whether everyone in the group will contribute part of their
2 earnings to provide needy families with a small flock of chickens. Each dollar per person will
3 provide an additional needy family with a flock of chickens.
4

5 ***Publicly-provided private good.*** Although careful attention was paid to make the public
6 and private good treatments as similar as possible, they differ in two important dimensions: the
7 elicitation mechanism (BDM vs. referendum) and the stage two good (coffee mug vs. Heifer
8 International). In the second stage, a vote in support of a group contribution to Heifer
9 International imposes costs on all group members and also provides a public good. The decision
10 to purchase a mug, however, is a private transaction that does not impose a cost on other group
11 members and the benefits also accrue only to the individual. To determine whether any
12 differences in behavior between the public and private good contexts are attributable to the
13 elicitation mechanism or the good itself, we also conducted a pair of publicly-provided private
14 good treatments.

15 The first stage was the same as that for the public good with a slightly different frame.
16 Subjects were asked to vote in a referendum (public provision) to determine whether everyone in
17 the group would buy a token (private good) at a randomly selected price. The pair of induced
18 values was the same as the other two treatments. In the second stage, subjects were asked to vote
19 in a referendum to determine whether everyone would buy a coffee mug. If a majority of
20 subjects voted yes to a randomly selected price, then each person would have to buy a mug
21 regardless of how she voted.
22

23 **4. Results**

24 ***Induced values.*** We begin using the stage one results to test whether hypothetical bias exists
25 with induced values. Because bids were constrained to whole dollar increments, a demand-

1 revealing individual with an induced value of \$4.50 would indicate a maximum WTP of \$4.00.
2 Similarly, when induced values are \$11.50, the maximum WTP should be \$11.00. Indeed, the
3 median WTP in both the real and hypothetical treatments conform to predictions for all three
4 good types and both payment amounts. Table 4 shows that with induced values of \$4.50, mean
5 hypothetical values are slightly higher than real payments, but median values were identical.
6 Mann-Whitney-Wilcoxon rank-sum (Wilcoxon 1945; Mann and Whitney 1947) tests indicate
7 that this between-subject difference is not significant for either the private or public goods.⁷ For
8 the publicly-provided private good, the difference is weakly significant ($p=0.08$); this result is
9 largely driven by four individuals in the hypothetical treatment who indicated a \$5 WTP which
10 would result in a \$0.50 loss. It is possible that earnings less than \$1 are not salient which could
11 help explain this outcome (Taylor, *et al.* 2001). Vossler and McKee 2006 use a similar threshold
12 to distinguish between small and large errors. When induced values are \$11.50, mean
13 hypothetical values are slightly lower than their real counterparts in two of the three treatments,
14 but again these differences are not significant. Hence, we cannot reject the hypothesis of no
15 hypothetical bias in five of the six comparisons of induced values (two induced values \times three
16 good types), and there is one case in which the null hypothesis is rejected, largely due to four
17 subjects willing to incur a 50-cent loss. This result is consistent with those of both Taylor, *et al.*
18 2001 and Vossler and McKee 2006 in a public goods context, and with Cherry *et al.* 2004 for
19 private goods.

20 Another measure for comparing bidding behavior between real and hypothetical payment
21 conditions is the mean absolute deviation (MAD) of the observed bid from the induced value
22 (Table 5). The MAD for subjects who fully reveals demand would be \$0.50 (resulting from a

⁷ Unless otherwise specified, all hypothesis tests conducted using Mann-Whitney-Wilcoxon rank-sum tests.

1 \$4.00 bid with a \$4.50 induced value or an \$11.00 bid with an \$11.50 induced value). Note that
2 with this approach, a bid that incurs a \$0.50 loss is treated identical to a demand revealing bid.
3 Using this metric, there is no statistically significant difference between real and hypothetical
4 bids for any of the comparisons in Table 5.

5 An advantage of the induced value experiments is that we can observe whether behavior
6 is consistent with these values, whereas with homegrown values we must *assume* that behavior
7 in the real payment treatments reflects the individual's true WTP. With the stage one induced
8 value data, we can identify misvotes or decision errors, *i.e.*, those instances in which individuals
9 are not behaving in a manner consistent with truthful demand revelation. In the private good
10 treatments, Table 6 shows that there were no misvotes in the real payment condition—which leaves
11 little doubt that subjects fully understood the instructions, the elicitation mechanism and how to
12 accurately state their values. Similarly, with hypothetical payments, only 3 of 56 subjects (5%)
13 misvoted and a Fisher exact test indicates that there is no statistically significant difference
14 ($p=0.25$) between misvotes in these hypothetical and real treatments.

15 This result of near-perfect bidding behavior using a BDM to elicit values for a private
16 good provides strong support for the demand revealing properties of the mechanism. With real
17 payments, Irwin *et al.* (1998) found that 62% of their BDM bids were optimal, and there was no
18 significant difference between bids and induced values; they did not have a hypothetical payment
19 condition for comparison. The no outside option treatments in Cherry *et al.* (2004) do allow for
20 a comparison of hypothetical and real payments in an induced value experiment for a private
21 good, but they use a standard second-price auction not a BDM.⁸ When averaged over the course
22 of 10 rounds, they do observe hypothetical bias: hypothetical bids are 38.5% higher than real

⁸ Shogren *et al.* (2001) note that second price auctions are accurate in aggregate, but are susceptible to insincere bidding by off-margin traders.

1 bids on average. Much of this is attributable to inflated hypothetical bids in later rounds, possibly
2 due to boredom. Their first two rounds represent a closer parallel with our experimental design.
3 Somewhat surprisingly, they report substantial *underbidding* in both real and payment conditions
4 for the first two rounds. Although real bids were 12.8% *higher* than hypothetical bids, as in our
5 study this difference was not statistically significant.⁹

6 When we introduced the referendum, however, the rate of misvotes increased. In our
7 public good referendum, we observe 19% and 12% misvotes in the real and hypothetical
8 treatments, respectively, and a Fisher exact test indicates that there is no statistically significant
9 difference between these ($p=0.44$). This result is consistent with those of Taylor *et al.* who
10 observe about a 16% misvote rate in both treatments. In the publicly-provided private good
11 treatment, hypothetical misvotes (34%) are significantly higher than in the real payment scenario
12 (12%). Much of this difference in the rate of misvotes between the real and hypothetical payment
13 scenarios is due to subjects voting yes to hypothetical outcomes that would yield a \$0.50 loss.
14 Both Taylor *et al.* and Vossler and McKee suggest that such small losses may not be salient;
15 using a “loose” definition of misvotes (which does not classify a vote leading to a 50 cent loss as
16 misvote) we find that the difference between real and hypothetical payments for the publicly
17 provided private good is much smaller and only weakly significant ($p=0.10$). As in Taylor *et al.*,
18 we find that although there are some misvotes at the individual level, those who overbid and
19 underbid tend to cancel each other out so aggregate results are approximately demand revealing
20 with both the real and hypothetical payments. However, this also raises questions about the
21 conditions under which these voting errors in a referendum might be asymmetric, possibly
22 leading to a bias in some referenda.

⁹ Thanks to Todd Cherry for sharing their data.

1 We find it particularly noteworthy that bidding behavior with private goods is near-
2 perfect, but when the BDM-style referendum is introduced some misvoting occurs. This raises
3 the question as to why misvoting arises more frequently with a group decision via referenda than
4 with a purely private transaction. Our design does not permit a testing of possible causes for this
5 outcome and we can only provide some conjectures about this. We note the rate of misvotes
6 observed is consistent with that reported by both Taylor *et al.* and Vossler and McKee. While
7 we cannot rule out the possibility that misvotes in the referendum are due to confusion or
8 misunderstanding, we feel that this is unlikely, particularly in light of the near-perfect outcomes
9 with the private good BDM mechanism which was structured to be as close as possible to the
10 BDM-style referendum we used. More likely, these misvotes are not random errors, but rather
11 indicate that a referendum is highly demand revealing, but not perfectly so. Vossler and McKee
12 suggest that some misvoting could be due to other-regarding behavior, in particular a willingness
13 to incur a small loss if it is in the best interest of the group. They conclude that this effect is
14 likely to be small and we tend to agree. Without information about the distribution of others'
15 induced values, subjects have no way of knowing what is in the group interest. Another possible
16 explanation is a misguided attempt at strategic behavior or coordination. The private good BDM
17 is a purely individual decision making exercise with no potential for strategic interactions or
18 other-regarding behavior. However, the referendum does entail a group decision and subjects
19 could conceivably make errors in recognizing that this mechanism is also incentive compatible.

20 ***Homegrown values.*** Whereas we observe no hypothetical bias with induced values, when
21 we elicit homegrown values this bias does emerge. The bottom of Table 4 shows that
22 hypothetical bias is present for all three types of good. Median WTP is \$3 to \$4 higher in the
23 three hypothetical payment treatments than their real payment counterparts, and these differences

1 are statistically significant. Thus, the same subjects who exhibited no hypothetical bias in stage
2 one with induced values, do exhibit hypothetical bias when asked to value a specific good in
3 stage two. These results clearly support Taylor *et al.*'s conjecture that hypothetical bias is not a
4 value elicitation problem in aggregate, but rather a value formation problem.

5 An unexpected result with the homegrown values is the higher willingness-to-pay for the
6 exact same coffee mug in the private good treatment when compared to the publicly-provided
7 context. The main difference between the two is the elicitation procedure: individual purchase
8 via an individual decision task versus a referendum requiring all group members to buy a mug.
9 The difference is significant for real payments ($p=0.08$), but not with hypothetical payments
10 ($p=0.22$). If anything, we anticipated a lower willingness-to-pay in the referendum because
11 participants might be reluctant to require others to involuntarily purchase a private good.

12

13 **5. Conclusions**

14 Induced valuation respondents are confronted with known values for the commodity being
15 examined. Hypothetical bias is therefore unlikely with induced values, but it could arise; in a
16 post-experiment survey some individuals in our real payment treatments with referenda
17 expressed concern about forcing others to pay. On the other hand, homegrown values
18 formulated during the CV exercise are likely to be quite uncertain and the potential for
19 hypothetical bias is much greater.

20 Although few studies have examined the process by which individuals make decisions in
21 the CV framework, Schkade and Payne's (1994) verbal protocol analysis suggests that
22 respondents in a hypothetical homegrown valuation focus on a variety of considerations. Among
23 these are the notion of paying ones' fair share, consideration of previous charitable contributions,

1 and a concern for “how the world should be”, all largely detached from the context of money.
2 But, when asked to make actual payment, respondents appear to focus on monetary
3 considerations, and this can give rise to hypothetical bias.

4 Of course, in homegrown situations many other factors, like free riding, may contribute
5 to this problem. And, in some cases, hypothetical bias may even be seen as a general cognitive
6 response; people like to think of themselves as being generous and this is costless when payment
7 is not required. Even our everyday speech hints at hypothetical bias—“actions speak louder than
8 words,” “practice what you preach”, and the “road to hell is paved with good intentions.”

9 To better understand and to correct for hypothetical bias, a necessary first step is to be
10 able to turn it off and on in the laboratory which is the major contribution of this paper. We
11 observe that hypothetical bias at the aggregate level emerges during the process of homegrown
12 value formation for private, publicly-provided private, and public goods. The presence of
13 individual decision errors in our referenda using induced values, but not in the BDM, suggests
14 that although value formation is a major factor in hypothetical bias, the elicitation mechanism
15 may also play a secondary role, at least at the individual level.

16

1 **References**

- 2 Aadland, D. and Caplan, A. J. 2003. Willingness to pay for curbside recycling with detection and
3 mitigation of hypothetical bias. *American Journal of Agricultural Economics*, 85(2), 492-
4 502.
- 5 Aadland, D. and Caplan, A. J. 2006. Cheap talk reconsidered: Evidence from CVM. *Journal of*
6 *Economics and Behavioral Organization*, 60(4), 562-578.
- 7 Becker, G., DeGroot, M., and Marschak, J. 1964. Measuring utility by a single-response
8 sequential method. *Behavioral Science*, 9(3), 226-232.
- 9 Blumenschein, K., Johannesson, M., Blomquist, G. C., Liljas, B. and O'Conor, R. M. 1998.
10 al results on expressed certainty and hypothetical bias in contingent valuation. *Southern*
11 *Economic Journal*, 65(1), 169-77.
- 12 Brown, T. C., Ajzen, I. and Hrubes, D. 2003. Further tests of entreaties to avoid hypothetical bias
13 in referendum contingent valuation. *Journal of Environmental Economics and Management*,
14 46(2), 353-361.
- 15 Burton, A. C., Carson, K. S., Chilton, S. M., Hutchinson, W. G. 2003. An experimental
16 investigation of explanations for inconsistencies in responses to second offers in double
17 referenda. *Journal of Environmental Economics and Management*, 46(3), 472-489.
- 18 Champ, P. A., R. C. Bishop, T. C. Brown, and D. W. McCollum. 1997. Using Donation
19 Mechanisms to Value Nonuse Benefits from Public Goods. *Journal of Environmental*
20 *Economics and Management*, 33(2): 151-162.
- 21 Champ, P. A. and Bishop, R. C. 2001. Donation payment mechanisms and contingent valuation:
22 an empirical study of hypothetical bias. *Environmental and Resource Economics*, 19(4), 383-
23 402.
- 24 Cherry, T. L., Frykblom, P., Shogren, J., List, J. and Williams, M. 2004. Laboratory testbeds and
25 nonmarket valuation: the case of bidding behavior in a second price auction with an outside
26 option. *Environmental and Resource Economics*, 29(3), 285-294.
- 27 Cummings, R. G., and Taylor, L. O. 1999. Unbiased value estimates for environmental goods: a
28 cheap talk design for the contingent valuation method. *American Economic Review*, 89(3),
29 649 - 665.
- 30 Ethier, R. G., G. L. Poe, W. D. Schulze, and J. Clark. 2000. A Comparison of Hypothetical
31 Phone and Mail Contingent Valuation Responses for Green-Pricing Electricity Programs.
32 *Land Economics*. 76(1): 54-67.
- 33 Gregory, R., S. Lichtenstein, T. C. Brown, G. L. Peterson, and P. Slovic. 1995. How Precise Are
34 Monetary Representations of Environmental Improvements? *Land Economics*. 71(4): 462-73.
- 35 Gregory, R., and P. Slovic. 1997. A Constructive Approach to Environmental Valuation.
36 *Ecological Economics*. 21(3): 175-181.
- 37 Harrison, G.W., R.M. Harstad, and E.E. Rutström. 2004. Experimental Methods and Elicitation
38 of Values. *Experimental Economics*. 7(2): 123-140

- 1 Harrison, G.W., and E.E. Rutström. 2008. Experimental Evidence on the Existence of
2 Hypothetical Bias in Value Elicitation Methods. In: Plott, C.R. and Smith, V.L. (eds).
3 *Handbook of Experimental Economics Results*, volume 1. New York: North-Holland.
- 4 Horowitz, J.K. 2006. The Becker-DeGroot-Marschak mechanism is not necessarily incentive
5 compatible, even for non-random goods. *Economics Letters*. 93:6-11.
- 6 Horowitz, J.K., McConnell, K.E. and J.J. Murphy. Forthcoming. Behavioral Foundations of
7 Environmental Economics and Valuation. In: List, J.A. and Price, M.F. (eds). *Handbook on*
8 *Experimental Economics and the Environment*. Northampton, MA: Edward Elgar.
- 9 Irwin, J.R., G.H. McClelland, M. McKee, W.D.Schulze, N.E. Norden. 1998. Payoff Dominance
10 vs. Cognitive Transparency in Decision Making. *Economic Inquiry*. 36(2):272-285.
- 11 Johannesson, M., B. Liljas, and P. Johannsson. 1998. An Experimental Comparison of
12 Dichotomous Choice Contingent Valuation Questions and Real Purchase Decisions. *Applied*
13 *Economics*. 30: 643-647.
- 14 Li, C.-Z., and L. Mattsson. 1995. Discrete Choice under Preference Uncertainty: An Improved
15 Structural Model for Contingent Valuation. *Journal of Environmental Economics and*
16 *Management*. 28(2): 256-269.
- 17 List, J. A. 2001. Do explicit warnings eliminate the hypothetical bias in elicitation procedures?
18 evidence from field auctions for sports cards. *American Economic Review*, 91(5), 1498-1507.
- 19 List, J. A. and Gallet, C. 2001. What experimental protocol influence disparities between actual
20 and hypothetical stated values? *Environmental and Resource Economics*, 20(3), 241-254.
- 21 Lusk, J.L. 2003. Willingness-to-pay for Golden Rice. *American Journal of Agricultural*
22 *Economics*. 85(4): 840-856.
- 23 Lusk, J.L. and M. Rousu. 2006. Market price endogeneity and accuracy of value elicitation
24 mechanisms. In: List, J.A. (ed). *Using Experimental Methods in Environmental and Resource*
25 *Economics*. Northampton, MA: Edward Elgar.
- 26 Mann, H. B., and D. R. Whitney. 1947. On a test whether one of two random variables is
27 stochastically larger than the other. *Annals of Mathematical Statistics*. 18: 50-60.
- 28 Mitani, Y. and Flores, N.E. 2009. Demand Revelation, Hypothetical Bias, and Threshold Public
29 Goods Provision. *Environmental and Resource Economics*. 44(2):231-243.
- 30 Murphy, J. J. and Stevens, T. H. 2004. Contingent valuation, hypothetical bias and experimental
31 economics. *Agricultural and Resource Economics Review*, 33(2), 182-192.
- 32 Murphy, J. J., Stevens, T. H., Allen, P. G. and Weatherhead, D. 2005a. A Meta-analysis of
33 hypothetical bias in stated preference valuation. *Environmental and Resource Economics*,
34 30(3), 313-325.
- 35 Murphy, J. J., Stevens, T. H., and Weatherhead, D. 2005b. Is cheap talk effective at eliminating
36 hypothetical bias in a provision point mechanism? *Environmental and Resource Economics*,
37 30(3), 327-343.
- 38 Noussair, C., Robin, S. and B. Ruffieux. 2004. Revealing consumers' willingness-to-pay: A
39 comparison of the BDM mechanism and the Vickrey auction. *Journal of Economic*
40 *Psychology*. 25:725-741.

1 Opaluch, J.J., Segerson, K., 1989. Rational roots of 'irrational' behavior: new theories of
2 economic decision-making. *Northeastern Journal of Agricultural and Resource Economics*
3 18(2):81-95.

4 Plott, C.R. and K. Zeiler. 2005. The willingness to pay–willingness to accept gap, the
5 “endowment effect,” subject misconceptions, and the experimental procedures for eliciting
6 evaluations. *American Economic Review*, 95(3):530-545.

7 Poe, G. L., Clark, J. E., Rondeau, D. and Schulze, W. D. 2002. Provision point mechanisms and
8 field validity tests of contingent valuation. *Environmental and Resource Economics*,
9 23(1):105-131.

10 Polomé, P. 2003. Experimental Evidence on Deliberate Misrepresentation in Referendum
11 Contingent Valuation. *Journal of Economic Behavior and Organization*. 52 (3): 387-401.

12 Schkade, D., and J.W. Payne. 1994. How People Respond to Contingent Valuation Questions: A
13 Verbal Protocol Analysis of Willingness to Pay for an Environmental Regulation. *Journal of*
14 *Environmental Economics and Management*. 26: 88-109.

15 Shogren, J.F., Margolis, M., Koo, C. and J.A. List. 2001. A Random *n*th-Price Auction. *Journal*
16 *of Economic Behavior and Organization*. 46(4):409-421.

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

19 Vossler, C. A. and McKee, M. 2006. Induced-Value Tests of Contingent Valuation Elicitation
20 Mechanisms. *Environmental and Resource Economics*, 35(2):137-168.

21 Wang, H. 1997. Treatment of 'Don't-Know' Responses in Contingent Valuation Surveys: A
22 Random Valuation Model. *Journal of Environmental Economics and Management*. 32(2):
23 219-232.

24 Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics*. 1: 80-83.
25

1 **Table 1. Experimental Design**

Type of Good	Payment Type (number of subjects)		Homegrown Value Good	Elicitation Mechanism
	Real	Hypothetical		
Private good	25	28	Coffee mug	BDM
Publicly-provided private good	29	29	Coffee mug	Referendum
Public Good	29	29	Heifer International	Referendum

2

3

1 **Table 2. Sequence of events within a treatment**

	Round	Value Type	Good offered
Practice	P1	Induced	Token value \$6.25
	P2	Induced	Token value \$11.60
<i>Practice results announced</i>			
Stage 1	1	Induced	Token value \$4.50
	2	Induced	Token value \$11.50
Stage 2	3	Homegrown	Coffee mug or Heifer International
<i>Stage 1 and 2 results announced</i>			

2

3

1 **Table 3. Example Purchase Offer Slip in the Induced Value Private Good Real Payment**
 2 **Treatment**

3
 4

Your Token Value is **\$4.50**.

Possible Prices	Indicate whether you are willing to buy a token at each possible price			
		If you <u>do</u> buy a token at this price, you will earn...		If you <u>do not</u> buy a token at this price, you will earn...
\$1	<input type="checkbox"/> Yes	\$3.50	<input type="checkbox"/> No	\$0
\$2	<input type="checkbox"/> Yes	\$2.50	<input type="checkbox"/> No	\$0
⋮	⋮	⋮	⋮	⋮
\$14	<input type="checkbox"/> Yes	– \$9.50	<input type="checkbox"/> No	\$0
\$15	<input type="checkbox"/> Yes	– \$10.50	<input type="checkbox"/> No	\$0

5
 6
 7

Please remember that your Purchase Offers must be consistent.

1 **Table 4. Mean and Median Willingness-to-Pay**
 2

Type of Good	Mean		Median		Wilcoxon (p-value) ^a
	Real	Hypothetical	Real	Hypothetical	
\$4.50 Induced Value					
Private	4.00 (0)	4.29 (0.25)	4.00	4.00	0.18
Public Private	3.93 (0.05)	4.34 (0.21)	4.00	4.00	0.08 *
Public	3.83 (0.13)	4.24 (0.18)	4.00	4.00	0.15
\$11.50 Induced Value					
Private	11.00 (0)	10.86 (0.14)	11.00	11.00	0.34
Public Private	10.76 (0.11)	10.21 (0.40)	11.00	11.00	0.57
Public	10.59 (0.31)	10.93 (0.23)	11.00	11.00	0.92
Homegrown Value					
Private	2.08 (0.45)	4.43 (0.68)	1.00	5.00	0.02 ***
Public Private	3.41 (0.56)	5.55 (0.62)	3.00	6.00	0.01 ***
Public	5.72 (0.86)	8.17 (0.94)	5.00	8.00	0.07 *

3 ^a Results of a Mann-Whitney-Wilcoxon test of the null hypothesis that the real and hypothetical values were
 4 drawn from the same distribution. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$. Standard errors in parentheses.
 5

1 **Table 5. Mean Absolute Deviations of Bids from Induced Values**
 2

Type of Good	Mean Absolute Deviation ^a		Wilcoxon (p-value) ^b
	Real	Hypothetical	
\$4.50 Induced Value			
Private	0.50 (0.00)	0.71 (0.21)	0.34
Public Private	0.57 (0.05)	0.81 (0.15)	0.13
Public	0.71 (0.13)	0.71 (0.13)	1.00
\$11.50 Induced Value			
Private	0.50 (0.00)	0.64 (0.14)	0.34
Public Private	0.74 (0.11)	1.50 (0.37)	0.16
Public	1.02 (0.29)	0.81 (0.20)	0.66

3 ^a Mean absolute deviation equals the induced value minus the observed bid. Standard errors in parentheses. If a
 4 subject submits a perfectly demand revealing bid, then this value is 0.50.
 5 ^b Results of a Mann-Whitney-Wilcoxon test of the null hypothesis that the real and hypothetical values were
 6 drawn from the same distribution. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

7
 8

1 **Table 6. Percent of Misvotes with Induced Values (Stage 1)**
 2

Percent Misvotes (both induced values combined)			
Type of Good	Real	Hypothetical	Fisher exact ^a
Strict Definition ^b			
Private	0%	5%	0.25
Public Private	12%	34%	0.01 ***
Public	19%	12%	0.44
Loose Definition ^c			
Private	0%	4%	0.50
Public Private	12%	26%	0.10 *
Public	12%	10%	1.00

3 * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

4 ^a p-values from a Fisher exact test comparing the rate of misvotes
 5 in the Real and Hypothetical treatments.

6 ^b A misvote occurs if a respondent indicates a WTP other than \$4 with a \$4.50 induced value, or a WTP other than
 7 \$11 with an \$11.50 induced value .

8 ^c A misvote occurs if a respondent indicates a WTP other than \$4 or \$5 with a \$4.50 induced value, or a WTP
 9 other than \$11 or \$12 with an \$11.50 induced value .