



University of Massachusetts Amherst

Department of Resource Economics

Working Paper No. 2004-9

<http://www.umass.edu/resec/workingpapers>

Hypothetical Bias in Dichotomous Choice Contingent Valuation Studies

Michael Ash¹, James J. Murphy², and Thomas H. Stevens³

Abstract:

This paper uses a meta-analysis to explore the relationship between hypothetical bias and the price respondents are asked to pay. For public goods, the results clearly indicate a difference in the price elasticity between hypothetical and actual payment conditions. Since the bias increases for larger dollar amounts, any simple guidelines, such as NOAA's "divide by two" rule of thumb, could be misleading. Future attempts to calibrate contingent valuation responses should reflect this price sensitivity.

Keywords: contingent valuation; experiments; hypothetical bias; meta-analysis; stated preference

JEL Classification: C9, Q26, Q28, H41

¹ Michael Ash, Department of Economics
University of Massachusetts, Thompson Hall
Amherst, MA 01003
E: mash@econs.umass.edu P: 413-545-6329

² James J. Murphy, Department of Resource Economics
University of Massachusetts, 219A Stockbridge Hall, 80 Campus Center Way
Amherst, MA 01003
E: murphy@resecon.umass.edu P: 413-545-5716 F: 413-545-5853

³ Thomas H. Stevens, Department of Resource Economics
University of Massachusetts, 216 Stockbridge Hall, 80 Campus Center Way
Amherst, MA 01003
E: tstevens@resecon.umass.edu P: 413-545-5714 F: 413-545-5853

April 2005

Hypothetical Bias in Dichotomous Choice Contingent Valuation Studies

MICHAEL ASH
Assistant Professor
Department of Economics &
Center for Public Policy and Administration
University of Massachusetts-Amherst

JAMES J. MURPHY*
Assistant Professor
Department of Resource Economics &
Center for Public Policy and Administration
University of Massachusetts-Amherst

THOMAS H. STEVENS
Professor
Department of Resource Economics
University of Massachusetts-Amherst

Correspondence to: James J. Murphy
Department of Resource Economics
219A Stockbridge Hall
80 Campus Center Way
University of Massachusetts-Amherst
Amherst, MA 01003, USA.
Phone: (413) 545-5716
Fax: (413) 545-5853
E-mail: murphy@resecon.umass.edu

* Corresponding author.

Hypothetical Bias in Dichotomous Choice Contingent Valuation Studies

Abstract

This paper uses a meta-analysis to explore the relationship between hypothetical bias and the amount respondents are asked to pay. For public goods, the results clearly indicate a difference in the price elasticity between hypothetical and actual payment conditions. Since the bias increases for larger dollar amounts, any simple guidelines, such as NOAA's "divide by two" rule of thumb, could be misleading. Future attempts to estimate hypothetical bias or to calibrate contingent valuation responses should reflect this price sensitivity. For private goods, however, hypothetical bias is constant across payment amounts.

Keywords: contingent valuation; experiments; hypothetical bias; meta-analysis; stated preference

JEL Classification: C9, Q26, Q28, H41

Hypothetical Bias in Dichotomous Choice Contingent Valuation Studies

Introduction

Contingent valuation (CV) surveys request hypothetical payment for the hypothetical provision of a non-market good. There is substantial evidence that these hypothetical values exceed actual payments (List and Gallet 2001; Murphy, *et al.* 2005; Harrison and Rutström, forthcoming). In recognition of this hypothetical bias problem, the NOAA guidelines state that CV value estimates should be halved unless responses can be calibrated (Arrow, *et al.* 1993). This spawned an emphasis on developing calibration techniques to correct for hypothetical bias. Some attempt to elicit unbiased responses, for example by using a cheap talk script (Cummings and Taylor 1999). Others try to calibrate biased responses using statistical bias functions (Blackburn, *et al.* 1994; Mansfield 1998; Hofler and List 2004) or uncertainty adjustments (Champ and Bishop 2001; Poe, *et al.* 2002).

In a synthesis of the hypothetical bias literature, Murphy and Stevens 2004 note that although there has been some success calibrating responses, surprisingly little research directly tests hypotheses about why these techniques work, or more generally on the underlying causes of the bias. They note that until the factors leading to hypothetical bias are better understood, calibration procedures to mitigate this effect may not be robust. For example, the effectiveness of the cheap talk methods appears sensitive to the length of the script. Cummings, *et al.* 1995 report that a short script was not effective at eliminating hypothetical bias, but a longer script similar to Cummings and Taylor 1999 was effective. List 2001 finds that the long script did not reduce hypothetical bias with experienced card dealers, but was effective with inexperienced participants, and Brown, *et al.* 2003 found that the long cheap talk script was successful, but only

for high payment amounts. Similarly, the level of certainty that accurately calibrates responses varies across studies.

In the absence of a theory to explain hypothetical bias, it is worth exploring existing studies to determine whether there are any empirical regularities that may help us better understand the conditions under which hypothetical bias arises, the extent to which the bias is sensitive to different factors, and circumstances under which calibration techniques are effective. Ultimately, identifying these relationships may lead to a better understanding of the nature of hypothetical bias.

Along these lines, both List and Gallet 2001, and Murphy, *et al.* 2005 use meta-analyses to test whether particular factors can explain differences in responses to real and hypothetical payments. These studies confirm the hypothetical bias problem and show that a key factor in explaining the bias is whether the good is public or private. List and Gallet 2001 found that willingness-to-pay studies have a lower bias than those that estimate willingness-to-accept, and that to some extent this bias may vary across elicitation mechanisms. Murphy, *et al.* 2005 note that the number of willingness-to-accept studies is quite small and List and Gallet's result is sensitive to the outcomes of a single study. Similarly, many elicitation mechanisms in List and Gallet have only a single study so their conclusions about elicitation mechanisms should be interpreted with caution. Because of the lack of variability in the types of elicitation mechanisms, Murphy, *et al.* 2005 group elicitation mechanisms based on whether they are choice-based (e.g., dichotomous choice, referendum, etc.), and find that these mechanisms yield less hypothetical bias. They also find some evidence that the use of student subjects may be a source of bias, but since this variable is highly correlated with group experimental settings, they refrain from drawing any firm conclusions.

However, a potentially important factor in explaining hypothetical bias has received little attention—the amount an individual is asked to pay. Since hypothetical payments lack a consequential economic commitment, we hypothesize that these responses would be relatively insensitive to changes in prices. Perhaps one of the most intuitively appealing explanations for hypothetical bias is that if the respondent has a positive value for a public good, and if her response to the valuation question may increase the likelihood of the good’s provision at little or no cost to her, then it makes sense for her to report an inflated value (Bohm 1984; Harrison and Rutström). It is also possible that individuals are merely expressing a positive attitude for the good without necessarily agreeing to contribute towards its provision (Champ and Bishop 2001). If people are trying to increase the likelihood of a good’s (costless) provision or expressing positive attitudes for the good, then it is likely that these responses to hypothetical CV questions would be less sensitive to changes in payment amounts. On the other hand, actual payments represent a salient financial outlay for the respondent, therefore we expect to see a higher price elasticity. If hypothetical responses are generally less price elastic than responses to real payment questions, as indeed our results suggest, then this would imply that hypothetical bias increases with the amount asked, and that any attempts to address the hypothetical bias problem ought to account for this difference.

Some studies vary the payment amount, but they usually do not directly address whether the magnitude of the bias is sensitive to this variable. Many other studies only ask about a single payment amount. A simple review of the literature does not yield an obvious relationship between the amount asked and hypothetical bias. For example, Brown, *et al.* 2003 found that hypothetical bias increases with the payment amount, and that the cheap talk script is more effective at reducing the bias for higher payment amounts. However, they also express

reluctance to generalize these results for two reasons. First, they note that their response rate to hypothetical payments is invariant to amount asked, which they acknowledge is unexpected because it is inconsistent with most other studies. Typically, willingness-to-pay is (and should be) downward sloping, even in hypothetical treatments. Second, they use a narrow range of payments (\$1-\$8) and therefore express reluctance to extrapolate their results to higher prices. Champ and Bishop 2001 use a much wider range of payments (\$24-\$288) but the ratio of hypothetical to actual payments is fairly constant over much of that range. On the other hand, the results in Blumenschein, *et al.* 2001 show the largest bias for the smallest payment amount.

Since no clear consensus emerges in the literature, we pool the data from these and other studies and use a meta-analysis to test whether there are any significant differences in the price elasticity of responses to hypothetical and actual payments. This paper is distinct from List and Gallet 2001, and Murphy, *et al.* 2005 in that we pool the *data* from these studies (specifically, the percent of “yes” responses to each payment amount), rather than the estimated results. Studies that report summary statistics and WTP estimation results, but not the data itself, were excluded from our analysis. Because of this approach, most of the studies in our analysis were not included in List and Gallet 2001 or Murphy, *et al.* 2005. Using response rates to dichotomous choice payment questions, we estimate demand functions for hypothetical and actual transactions, and find clear evidence that hypothetical bias increases with the amount asked for public goods. For private goods, however, it appears that hypothetical bias remains constant across payment amounts.

Data and Methods

A review of the literature initially identified 37 dichotomous choice (including referenda) contingent valuation experiments that reported both hypothetical and actual payments.

Dichotomous choice studies were chosen because this elicitation format is recommended by NOAA and because these were the only studies that tended to report both actual and hypothetical data. The following criteria were then used to identify the studies included in our analysis:

- (1) The hypothetical and actual values had to be elicited from the same dichotomous choice format. This avoids confounding effects from different elicitation mechanisms with hypothetical bias.
- (2) Willingness-to-accept observations were excluded because very few studies have produced both actual and hypothetical willingness-to-accept values.
- (3) Since we wanted unadjusted hypothetical responses, data derived from ex ante or ex post attempts to reduce hypothetical bias, like use of cheap talk or uncertainty adjustments, were excluded.

Our final data set consists of 99 observations from 15 studies of 16 distinct goods.¹ Each observation represents the demand for a good under either hypothetical or actual payment conditions. We used the percent of respondents (quantity) willing to pay each amount asked (price) as the outcome variable and a measure of demand. For example, the Blumenschein, *et al.* 2001 study of willingness to pay for asthma management contributed six observations to our analysis: the percent of respondents answering “yes” when payment for and receipt of the good

¹ Our log specification resulted in the loss of one observation with zero demand. The results were unaffected when we preserved the zero-demand observation by adding a small positive amount to demand for each observation before taking the log.

were hypothetical at \$15, \$40, and \$80; and the percent answering “yes” when payment and receipt were actual at the same three prices.

Using equation 1, we estimate log demand ($\ln PctYes$ is the natural log of the percent of participants responding “yes”) as a function of both log price ($\ln AmtAsked$ is the natural log of the amount asked) and the interaction of $\ln AmtAsked$ with a dummy variable (Hyp) that equals one for a hypothetical payment condition. In order to pool studies that involve different goods with differing underlying valuations, we estimate the demand model with fixed effects for each combination of good ($i=1, \dots, 16$) and payment condition ($j=1$ for hypothetical and 0 for actual). Since the design of each study involved experimentally assigned price offers, indexed by k , we do not face the problem of simultaneous determination of price and quantity. An estimate of the quantity-price relationship from observed quantity-price pairs directly measures the price-responsiveness of demand. Thus, we identify the price elasticity of demand from within-good, within-condition variation in price. Using this specification, we can distinguish between hypothetical and actual demand for the good at the offered price.

$$\ln PctYes_{ijk} = \alpha_{ij} + \beta \cdot \ln AmtAsked_{ijk} + \delta \cdot \ln AmtAsked_{ijk} \cdot Hyp_{ij} + \varepsilon_{ijk} \quad (1)$$

The log specification allows us to interpret the coefficients as elasticities and provides a direct test of our basic hypothesis: the price elasticity of demand will differ between hypothetical and actual payment scenarios. β is the price elasticity for actual payments, and $\beta + \delta$ is the price elasticity for hypothetical payments; both should be negative. The main coefficient of interest is δ : the difference in the price elasticity for hypothetical payments relative to actual. Since hypothetical payments are not salient, we expect that the percent “yes” responses will be less sensitive to changes in the amount asked. This implies that δ should be positive and statistically significant.

Estimates of Hypothetical and Actual Demand

There are some conceptual differences in the valuation of public and private goods that suggest it might be inappropriate to pool these data, and there is some empirical evidence supporting this assertion. List and Gallet 2001 find that hypothetical bias is less for private goods, and Murphy, *et al.* 2005 report similar results for some model specifications. Therefore, Table 1 reports estimates for equation 1 using only public goods (our primary focus), only private goods, and both goods.

Roughly half the experimental studies of non-market valuation techniques use private goods and, as with public goods, hypothetical bias is usually observed. Nevertheless, there are also some fundamental differences between these two types of goods that may influence the way people respond to valuation questions and the nature of hypothetical bias. For example, individuals have more experience valuing private goods, even if they are not familiar with the particular good in question. Therefore, they may be more comfortable with the process of determining their value for the good and perhaps less prone to error. Hence, even when payments are hypothetical, individuals may be able to provide reasonable estimates of private good WTP with little effort. On the other hand, people are rarely asked to value public goods, and therefore may have less experience in developing these values. When payments for a public good are inconsequential, people may have little incentive to invest the cognitive energy to determine this value.

Another key difference is the existence of a market price for private goods. Both Harrison, *et al.* 2002 and Murphy and Stevens 2004 hypothesize that responses to actual payment questions for private goods may be censored by the market price; however, in an induced value

setting, Cherry, *et al.* 2004 do not find any differences in bid shaving between actual and hypothetical treatments when subjects have the outside option of purchasing a perfect substitute. Whether this result also holds in a homegrown value setting is still unresolved. Finally, whereas subjects might perceive that positive responses could increase the likelihood of a public goods' provision or they are simply expressing a positive attitude towards the public good, these motivations are not likely to be an issue when valuing private goods.

As one would expect, the results in Table 1 show that demand for both public and private goods is downward sloping: $\beta < 0$ and statistically significant. In addition, since private goods tend to have more substitutes, one would also expect that these goods would be more responsive to changes in price than public goods. The results in Table 1 are consistent with this: the price elasticity for the actual purchase of private goods ($\beta = -1.29$) is greater in absolute value than that for public goods ($\beta = -0.49$). More importantly, the results in Table 1 strongly support our hypothesis of different price elasticities for public goods; specifically, $\delta = 0.33$ is positive and significant. For actual payments, the price elasticity is -0.49 . However, when payments are hypothetical, the demand elasticity is $-0.16 (= -0.49 + 0.33)$, which indicates that these "yes" responses are less sensitive to changes in amount asked.

Figure 1 illustrates this relationship for public goods. The figure shows within-good and within-treatment mean-deviations of log price (amount asked) and log quantity (percent yes) for the public goods experiments. The grand mean for actual and the grand mean for hypothetical treatments were added to center each demand curve at the predicted values for a generic public good in actual and hypothetical settings.

For larger dollar amounts, this elasticity difference can have a sizeable impact on bias estimates. Table 2 reports calculations of the estimated average bias for public goods with

different amounts asked. In the fixed effects model from equation (1), each combination of good and payment condition has a different intercept (α_{ij}). For the average bias calculations in Table 2, we use the average α_{ij} for each payment condition ($\alpha_{i0} = 4.45$; $\alpha_{i1} = 4.28$). The amounts asked were chosen based on the distribution of these values in our data. It turns out that at the median amount asked (\$9), the NOAA “divide by 2” rule yields a reasonably accurate estimate, but as the amounts increase, so does the bias. We emphasize these bias estimates are for illustrative purposes and should not be used to calibrate CV results because the individual fixed effects are not included in the calculation.

The evidence for a difference in elasticities with private goods is less compelling: although δ remains positive, the estimate is small (0.09) and no longer significant ($p=0.84$). Note that δ only reflects the difference in elasticity, and alone does not indicate the magnitude of the hypothetical bias. If hypothetical bias is constant across amount asked for good i (as is typically assumed), then the difference in the fixed effects, $\alpha_{i1} - \alpha_{i0}$, would be positive and significant, but δ would not be significant. Hence, although most of these private good studies observe hypothetical bias, the bias is insensitive to the amount asked. This result is consistent with the conjecture that people may be more comfortable valuing goods they commonly purchase and may be less prone to error at the margin (List and Gallet 2001).

Discussion

Our analysis produces two noteworthy results. First, there is strong evidence indicating that for public goods, which is the focus of CV studies, hypothetical bias is sensitive to the amount asked. Attempts to understand hypothetical bias or to calibrate responses ought to account for this effect. Interestingly, at least with respect to the set of dichotomous choice public good

studies examined here, hypothetical bias appears to be roughly consistent with the NOAA panel recommendation around the median amount asked. Second, whereas hypothetical payments are less sensitive to changes in price for public goods, there does not appear to be a difference in the price elasticity across hypothetical and actual payments for private goods. This could suggest that there are some fundamental differences in the nature of hypothetical bias between these two types of goods that warrant further investigation. Considering that roughly half the hypothetical bias studies use private goods, it is important to understand how this bias may differ across commodity types and opens the question about the conditions under which results with private goods can be generalized to public good contexts.

Since a formal theoretical basis for hypothetical bias has not yet been developed, the nature of hypothetical bias, at least in terms of magnitude and functional form, is largely an empirical issue. The analysis presented here indicates that for public goods the magnitude of hypothetical bias is sensitive to the amount asked, but that this result does not hold for private goods. This implies that simple guidelines, such as those developed by NOAA, could be misleading. Provision of local public goods such as water supply, schools, and open space often involve substantial cost to individuals. Our results suggest that application of CV in these situations may be quite problematic, particularly if individuals are asked to pay a significant amount. This problem may also arise in some large-scale development projects undertaken by international organizations such as the World Bank. That is, it may be inappropriate to use the traditional CV method for projects and programs that would require relatively large payments from each individual. Calibration techniques, such as cheap talk and uncertainty adjustments, may mitigate this bias. However, the interaction of the techniques with various payments amounts needs further study. We are aware of only a single study, Brown, *et al.* 2003, that

investigates this relationship; they found that cheap talk is more effective at higher payment amounts.

Further research on the relationship between alternative CV response formats, hypothetical bias, and payment amount may also prove fruitful. For example, “yea-saying,” wherein respondents tend to anchor on the posited bid amount, has often been associated with the dichotomous choice format used in all studies examined here. Yea-saying that is more prevalent in costless hypothetical scenarios as compared to real payment situations may be one factor associated with our results. Whether these findings are associated with other response formats remains an interesting but open question.

The processes used by individuals responding to CV questions and the factors responsible for hypothetical bias remain unknown and we believe that continued investigation of the behavior and underlying causes of hypothetical bias, particularly at relatively high payment levels, is an important and useful endeavor. More research focused on understanding individual motives underlying hypothetical bias over a wide range of payment amounts is clearly needed.

Table 1. Estimates of the price elasticity of hypothetical and actual demand ^a

Variable	Public Goods		Private Goods		All observations	
	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error
Constant (α_{ij})	Unique fixed effect for each combination of good \times payment condition ^b					
lnAmtAsked (β)	-0.49 ***	0.06	-1.29 ***	0.41	-0.67 ***	0.12
lnAmtAsked \times Hyp (δ)	0.33 ***	0.08	0.09	0.37	0.24	0.17
N	54		45		99	
F	43.9		18.66		21.08	
Prob>F	0.00		0.00		0.00	
R ²	0.72		0.56		0.39	

^a Dependent variable is the natural log of the percent of “yes” responses. *** denotes significant at 1%.

^b Coefficient estimates are available upon request.

Table 2. Estimated average hypothetical bias for public goods ^a

Amount Asked	Percentile	Bias ^b
\$1.00	Minimum	0.84
\$2.82	10%	1.19
\$4.23	25%	1.35
\$9.00	50%	1.74
\$18.75	75%	2.21
\$30.43	Mean	2.59
\$88.00	90%	3.67
\$288.00	Maximum	5.41

^a Bias is defined as the ratio of estimated hypothetical and actual percent of respondents willing to pay the amount asked. Bias equals one if there is no hypothetical bias and exceeds one if hypothetical bias is present.

^b Values were calculated using the public good coefficient estimates in Table 1. For the constant, we used the average fixed effect for each payment condition ($\alpha_{i0} = 4.45$; $\alpha_{i1} = 4.28$).

References

- Arrow, K., R. Solow, P. Portney, E. E. Leamer, R. Radner, and H. Schuman. 1993. "Report of the NOAA Panel on Contingent Valuation." *Federal Register*. 58(10): 4602-4614.
- Blackburn, M., G. W. Harrison, and E. E. Rutström. 1994. "Statistical Bias Functions and Informative Hypothetical Surveys." *American Journal of Agricultural Economics*. 76: 1084-1088.
- Blumenschein, K., M. Johannesson, K. K. Yohoyama, and P. R. Freeman. 2001. "Hypothetical versus Real Willingness to Pay in the Health Care Sector: Results from a Field Experiment." *Journal of Health Economics*. 20(3): 441-457.
- Bohm, P. 1984. "Revealing Demand for an Actual Public Good." *Journal of Public Economics*. 24(2): 135-51.
- Brown, T. C., I. Ajzen, and D. Hrubes. 2003. "Further Tests of Entreaties to Avoid Hypothetical Bias in Referendum Contingent Valuation." *Journal of Environmental Economics and Management*. 46(2): 353-361.
- Brown, T. C., I. Ajzen, and D. Hrubes. 2003. "Further Tests of Entreaties to Avoid Hypothetical Bias in Referendum Contingent Valuation." *Journal of Environmental Economics and Management*. 46: 353-361.
- Champ, P. A., and R. C. Bishop. 2001. "Donation Payment Mechanisms and Contingent Valuation: An Empirical Study of Hypothetical Bias." *Environmental and Resource Economics*. 19(4): 383-402.
- Cherry, T. L., P. Frykblom, J. F. Shogren, J. A. List, and M. B. Sullivan. 2004. "Laboratory Testbeds and Non-Market Valuation: The Case of Bidding Behavior in a Second-Price Auction with an Outside Option." *Environmental and Resource Economics*. 29(3): 285-294.
- Cummings, R. G., G. W. Harrison, and L. L. Osborne. 1995. *Can the Bias of Contingent Valuation Surveys Be Reduced?* Economics working paper. Columbia, SC: Division of Research, College of Business Administration, Univ. of South Carolina.
- Cummings, R. G., and L. O. Taylor. 1999. "Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method." *The American Economic Review*. 89(3): 649 - 665.
- Harrison, G. W., R. M. Harstad, and E. E. Ruström. 2002. *Experimental Methods and Elicitation of Values*. August.
- Harrison, G. W., and E. E. Rutström. forthcoming. "Experimental Evidence on the Existence of Hypothetical Bias in Value Elicitation Methods." In *Handbook of Results in Experimental Economics*. ed. C. Plott, and V. L. Smith. New York: Elsevier Science.
- Hofler, R. A., and J. List. 2004. "Valuation on the frontier: Calibrating Actual and Hypothetical Statements of Value." *American Journal of Agricultural Economics*. 86(1): 213-221.

- List, J. A. 2001. "Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions for Sports cards." *American Economic Review*. 91(5): 1498-1507.
- List, J. A., and C. Gallet. 2001. "What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values?" *Environmental and Resource Economics*. 20: 241-254.
- Mansfield, C. 1998. "A Consistent Method for Calibrating Contingent Valuation Survey Data." *Southern Economic Journal*. 64(3): 665-681.
- Murphy, J. J., and T. H. Stevens. 2004. "Contingent Valuation, Hypothetical Bias and Experimental Economics." *Agricultural and Resource Economics Review*. 33(2): 182-192.
- Murphy, J. J., T. H. Stevens, P. G. Allen, and D. Weatherhead. 2005. "A Meta-Analysis of Hypothetical Bias in Stated Preference Valuation." *Environmental and Resource Economics*. 30(3): 313-325.
- Poe, G. L., J. E. Clark, D. Rondeau, and W. D. Schulze. 2002. "Provision Point Mechanisms and Field Validity Tests of Contingent Valuation." *Environmental and Resource Economics*. 23: 105-131.