EXTERNAL VALIDITY OF CONTINGENT VALUATION: COMPARING HYPOTHETICAL AND ACTUAL PAYMENTS

Running Head: External Validity of Contingent Valuation

ABSTRACT

Whilst contingent valuation is increasingly used in economics to value benefits, questions remain concerning its external validity i.e. do hypothetical responses match actual responses? We present results from the first within sample field test. Whilst Hypothetical No are always an Actual No, Hypothetical Yes exceed Actual Yes responses. A constant rate of response reversals across bids/prices could suggest theoretically consistent option value responses. Certainty calibrations (verbal and numerical response scales) minimize hypothetical-actual discrepancies offering a useful solution. Helping respondents resolve uncertainty may reduce the discrepancy between hypothetical and actual payments, and thus lead to more accurate policy recommendations.

Word Count: 1933 (excluding references)
Tables: 5
Figures: 0

1.   INTRODUCTION

Contingent valuation (CV), willingness to pay elicited in a hypothetical survey, is used to value non-marketed goods. However, it has seen limited use at the health policy level, with cost-utility analysis (Quality Adjusted Life Years, QALYs) dominating. One concern with CV, which may explain its limited take-up, is the extent to which hypothetical responses reflect actual behaviour (Hausman, 2012). Economic research carried out in the laboratory and field, on home-grown and induced values, on public and private goods and on different elicitation mechanisms show that the proportion of hypothetical yes responses exceeds actual yes responses and that adjusting for certainty reduces disparities (List and Gallet, 2001; Little and Berrens 2004). However, to date, no study has used a within sample design, comparing, hypothetical and actual values for the same individual.[1] Within sample designs have the advantage that the same respondent answers the hypothetical and actual choice; thus, any differences in responses are less likely the result of sample selection, unobserved heterogeneity or econometric modelling (as may be the case for between sample studies and travel cost or hedonic pricing methods).

So far, six field studies have compared hypothetical and actual values in health economics. Blumenschein et al. (2001), Blumenschein et al (2008) and Blomquist et al. (2009) all used a between sample design to compare hypothetical and actual values, and found evidence of overestimation of WTP in a hypothetical context. However, these studies have the limitation that the difference between hypothetical and actual WTP may be explained by different samples receiving the two questions. Bhatia and Fox-Rushby (2003) and Bryan and Jowett (2010) used a within-subject design and

---

[1] Carlsson and Martinnsson (2001) used a within sample design to test the external validity of discrete choice experiments within an environmental setting in the context of (marginal) WTP.

found no evidence of a divergence, whilst Onwujekwe et al (2005), adopting a similar study design, found evidence of a difference. However, all three within sample studies compared hypothetical values to a fixed actual price purchase, thus limiting the applicability of their conclusions.

We take the literature forward by conducting the first within-subject study design that allows price offered in the actual transaction to exactly mimic what was offered in the hypothetical question. We also consider the use of certainty calibration to reduce divergence between hypothetical and actual responses.

2. STUDY DESIGN

There were three phases to the study:

2.1 Phase 1: First-round questionnaire

The questionnaire informed respondents that the Sa Kaeo Chief Medical Office was considering introducing a prepaid oral care plan (POCP) to improve oral health among preschool children. The POCP was described and WTP for a six-month enrolment elicited using a dichotomous choice question. The bid vector had 4 levels: 200, 250, 300 and 400 Thai Baht per 6 months. The lower limit was the actual cost of the plan (see Phase 3), while the range was informed by an open-ended pilot WTP study. Subjects were randomly allocated across bids.

Certainty of responses was collected using a verbal and numeric certainty scale. For the verbal scale respondents were asked how sure they were they would pay the specified amount, with five responses from "Yes, absolutely" to "No, absolutely not".

For the numeric certainty scale subjects were asked to rate the degree of certainty on a scale of 0 ("Absolutely uncertain") to 10 ("Absolutely certain").

First-round questionnaires were distributed over a weekend by health care workers and volunteers. They returned the following weekend to collect questionnaires. Respondents were asked to complete unanswered questions.

2.2 Phase 2: Second-round questionnaire

All subjects who completed the first questionnaire received a second questionnaire, distributed 1.5 weeks later. Respondent's names were identified in the first round questionnaire, and second round questionnaires were addressed to them. If that person was not at home, the recipient of the second round questionnaire was requested to consult that person for completion.

Subjects were informed that the Provincial Chief Medical Office (PCMO) had decided to introduce the POCP. They were reminded what the POCP involved, and told that the person who dropped off the second questionnaire would return within a few days to collect payment if the offer was accepted. Subjects were offered the POCP at the price it had been offered to them in the hypothetical survey and asked to indicate their decision and sign the questionnaire. All subjects who returned the first-round questionnaire received the second-round questionnaire (regardless of whether they answered Yes or No).

2.3 Phase 3: Premium collection

Within two days of the second-round questionnaires being distributed health care workers collected the money from those who indicated they wanted to purchase the POCP. In practice policyholders would be charged a flat rate premium of 200 Baht, set

at the unit cost plus 20% administration fees. Respondents who paid more than 200 Baht received a rebate the size of the difference. Upon payment respondents received insurance cover.

2.4 Testing for anchoring

To test whether the actual purchase decision was influenced by receiving the hypothetical questionnaire (and feeling committed to being consistent across survey's phases) the actual purchase decision of those answering the hypothetical question was compared with a sample receiving only a real choice. The questionnaire was distributed at the same time as the second-round questionnaire and was collected at the same time as the premium.

2.5 Sample and setting

The target population was all parents or guardians of pre-school children aged 3-6 years enrolled in the 3 private kindergartens in the Muang district of the Sa Kaeo province in Thailand. Of a target sample of 674 parents/guardians, 240 were randomly selected and assigned across bids in the hypothetical questionnaire. 200 individuals never exposed to the hypothetical questionnaire were also randomly assigned to one of the four-bids. The study was approved by the Ethical Review Committee of the Sa Kaeo Provincial Chief Medical Office.

3.   DATA ANALYSIS

Differences in the average proportions of Hypothetical and Actual Yes responses are calculated across bids and overall (significance tested through McNemar's chi-squared

tests). The same proportions were also compared after calibrating Hypothetical Yes responses through certainty (verbal and numeric scales). Utilising the within-individual nature of our design the coincidence of Hypothetical and Actual Yes responses by individual is also calculated. Average WTP values are computed through the non-parametric Turnbull estimator (Haab and McConnell, 2002). Finally, anchoring is tested by comparing response proportions to the proportions of the sample who only received the actual choice.

4. RESULTS

Of 240 questionnaires distributed, 146 were completed. Of these 146 respondents, 135 completed and returned the second-round questionnaire (56.25% response rate). Of the 200 Actual only questionnaires distributed, 122 were returned (61% response rate). Table I presents descriptive statistics of the Hypothetical-Actual and Actual only samples. The sample has a mean age of about 35 years old, with about 40% males. No significant differences are found across samples.

4.1 External validity

Table II reports the percentage of Hypothetical Yes and Actual Yes responses, both of which decrease as bids increase. The proportion of Hypothetical Yes responses consistently significantly exceeds Actual Yes responses for all bids, and overall (statistically significant at 5% level). Table III disentangles response patterns for Hypothetical followed by Actual response. Proportions of Yes-No responses vary little across bids, suggesting that the growing divergence between hypothetical and actual can be attributed to the dropping proportion of Actual Yes. With the exception of one

individual, all Hypothetical No remain an Actual No when offered the chance to purchase the service.

Columns 3 to 6 (Table II) presents certainty corrections. Coding as Yes only those who were absolutely certain (Absolutely Yes) eliminates statistical differences with Actual Yes responses. Coding as Yes only those who state Certainty 9-10 has similar results. In contrast, taking Certainty 8-10 or Certainty 10 as Yes responses either under- or over-estimates Actual Yes proportions (i.e. under- and over-compensating for the hypothetical nature of the question). Note, that no other individual characteristic explained any of part of the Hypothetical-Actual divergence.[2]

At the individual level (Table IV), only 39% of respondents remain consistent between the Hypothetical Yes and the Real Yes response, with the proportion rising to above 80% when we calibrate with certainty.

4.2 Willingness to pay

Similar patterns are found for WTP calculations (Table V). Hypothetical WTP is significantly greater than Actual WTP but no statistical differences are observed once calibrating through certainty. As before, coding as Yes only those with certainty of 10 appears too strict and underestimates WTP values.

4.2 Anchoring

Comparisons of the first and last columns of Tables II and V provide anchoring tests. Proportions of Actual Yes responses, as well as WTP values, in the Hypothetical-Actual and Actual Only samples are almost identical, suggesting minimal concerns for bias.

---

[2] Making the Hypothetical-Actual divergence a binary variable and a function of individual characteristics, none of the variables in Table 1 was significant apart from certainty. Results are available from the authors.

5.   DISCUSSION

We present the first within sample test of external validity of CV where the price offered in the real choice is the same as that offered in the hypothetical survey. While all Hypothetical No responses correspond to an Actual No the average percentage of Hypothetical Yes responses significantly exceeds Actual Yes responses. Whilst the proportion of Yes-No response patterns remains constant across bids, we observe the decreasing Yes-Yes proportions as bids increase being absorbed by increasing No-No proportions (as predicted by economic theory). Calibrating responses through certainty makes average Hypothetical and Actual Yes proportions and WTP values statistically indistinguishable. Our results partly confirm Blomquist et al. (2009) where numeric certainty values near 10 produced equivalence between hypothetical and actual decisions, whereas we further observe that certainty values of 10 could be exaggerating the correction, underestimating Yes proportions and WTP.

Conjecturing on the reasons behind hypothetical-real divergence, strategic answering (i.e. free-riding incentives) could partly explain results, though such explanation would require stronger evidence of decreasing degree-of-bias as the offered bid falls (a finding predicted by free-riding but not hypothetical bias). . Further, given our experimental context of a private insurance decision, free-riding is likely to be less of a concern. However, over-statement of purchase decisions is predicted by economic theory as over-statements of intentions are costless to the respondent and maximise future choice sets.[3] The fact that the POCP insurance was a one-off six-monthly plan should, in principle, help mitigate such behaviours. Yet, option value based responses are likely to be affected by subjects' response (un)certainty. In fact we observe a

---

[3] We thank a reviewer for this point.

positive relationship between degree of certainty and level of economic commitment; the POCP plan was only ever bought in the actual transaction by those stating a certainty above 7 (at any bid) and a certainty above 9 at the two highest bid levels. This systematic link was not observed with any other individual characteristic. Helping respondents resolve uncertainty may reduce discrepancies between hypothetical and actual purchases. Future research should attempt to control for this when investigating the impact of uncertainty, and indeed any other factors, on hypothetical purchases. Such research could even pursue the development of an economic-framework based on the interaction between uncertainty, option value and hypothetical behaviour.

A few remarks on the study. The lower level of the bid vector is determined by the true cost of the plan. Given the limited support for the left tail of the WTP distribution, WTP values are potentially inflated. However, given the purpose of the study is to test external validity the absolute WTP values are not of concern.

A short delay of two days between the second-round questionnaire and premium collection was chosen. This could cash-constrain the actual purchase decision, increasing the possibility of Yes-No responses. However, forewarning respondents that they should answer taking into account their money readily available and that the amount stated would need to be paid imminently should mediate such concerns. On the other hand, a longer delay could increase the risk of disclosure (i.e. respondents realising that others had been offered the good at a different price) increasing the chances of protest reactions manifesting themselves as hypothetical/real divergence. Comparable proportion of Yes-No responses across bids along with basic evidence

from a qualitative question on the reasons behind reversals would suggest that facing differential bids were not the drivers behind hypothetical-real discrepancies. [4]

In conclusion, this is the first within-subject study comparing stated and actual WTP values in a CV context. Whilst hypothetical and actual responses differ, adjusting for certainty produces promising results. Further research exploring reasons behind hypothetical/actual values divergence and adjustment mechanisms is necessary if CV studies are to inform health policy.

---

[4] Two weeks after the experiment a questionnaire debriefing respondents was circulated which also explored motivations behind response reversals. While, only 11 questionnaires were returned (thus limiting usefulness) none of the respondents cited differential bids as their reasoning.

REFERENCES

Bhatia, MR. and Fox-Rushby, JA. 2003. Validity of willingness to pay: Hypothetical versus actual payment, Applied Economics Letters, 10(12), 737- 740.

Blomquist, G., Blumenschein, K., and Johannesson, M. 2009. Eliciting willingness to pay without bias using follow-up certainty statements: Comparisons between Probably/Definitely and a 10-point certainty scale. Environmental & Resource Economics, 43(4), 473-502.

Blumenschein K, Johannesson M, Yokoyama KK, Freeman PR. 2001. Hypothetical versus real willingness to pay in the health care sector: results from a field experiment. Journal of Health Economics, 20, 441-457.

Blumenschein, K., Blomquist, G. C., Johannesson, M., Horn, N., and Freeman, P. 2008. Eliciting willingness to pay without bias: Evidence from a field experiment. Economic Journal, 118, 525, 114-137.

Bryan, S and Jowett, S. 2010. Hypothetical versus real preferences: results from an opportunistic field experiment. Health Economics, 19(12), 1502-9.

Carlsson, F. and Martinsson, P. 2001. Do Hypothetical and Actual Marginal Willingness to pay Differ in Choice Experiments? Journal of Environmental Economics and Management 41:179-192.

Haab, T. C. and McConnell, K. E. 2002. Valuing environmental and natural resources: The econometrics of non-market valuation. Cheltenham, UK: Edward Elgar.

Hausman, J. 2012. Contingent Valuation: From Dubious to Hopeless. Journal of Economic Perspectives, 26(4), 43-56.

List, J.A. and Gallet, C. A. 2001. What experimental protocol influence disparities between actual and hypothetical stated values? Environmental & Resource Economics 20: 241-254

Little, J. & Berrens, R. 2004. Explaining disparities between actually and hypothetical stated values: further investigation using meta-analysis. Economic Bulletin 3: 1-13.

Onwujekwe O, Hanson, K. Fox-Rushby J. 2005 Do divergences between stated and actual willingness to pay signify the existence of bias in contingent valuation surveys? Social Science and Medicine, 60(3):525-36.

ACKNOWLEDGMENTS

Table I
Descriptive statistics for samples [a, b]

	Hypothetical – Actual sample	Actual Only sample
Age	34.56	34.46
Male	42.42	37.19
Education (Primary/Junior high)	18.94	22.31
Occupation (Academic job)	14.07	13.93
Occupation (Civil servant)	32.59	40.16
Accommodation (Owner of house)	67.91	64.46
Income less than 15K Bhat	65.65	56.56
Family size	4.45	4.39
# Kids within HH under 18 years	1.88	1.96
Regularly see a dentist (Yes)	18.52	24.59
Saw dentist for last time 1-2 years ago	36.30	33.06
Oral concern – Very important	68.15	71.67
Did your child have a toothache in the past 2 years? (Yes)	40.74	44.26
Dental clinic is less than 10K from home	57.89	63.11
Spend 500Bhat or more on Entertainment (tape cassette, Video, CDs) per year	29.23	30.33

[a] Sample statistics indicate mean values for continuous variables and percentages for categorical ones.
[b] No significant difference at 5% when comparing the Hypothetical-Actual to Actual Only.

Table II

Average proportion of Yes responses across samples, bid levels and certainty calibrations

	Hypothetical – Actual sample						Actual Only sample
	Actual Yes	Hypothetical Yes	Absolutely Yes	Certainty 8-10	Certainty 9-10	Certainty 10	Actual Yes
200	12/36	20/36	8/36	15/36	10/36	5/36	10/32
	(33.3%)	(55.6%)*	(22.2%)	(41.7%)	(27.8%)	(13.9%)*	(31.3%)
250	7/33	16/33	5/33	9/33	5/33	3/33	8/34
	(21.2%)	(48.5%)*	(15.2%)	(27.3%)	(15.2%)	(9.1%)*	(23.5%)
300	4/35	15/35	5/35	8/35	5/35	4/35	3/29
	(11.4%)	(42.9%)*	(14.3%)	(22.9%)*	(14.3%)	(11.4%)	(10.3%)
400	2/31	10/31	3/31	6/31	3/31	3/31	1/27
	(6.5%)	(32.3%)*	(9.7%)	(19.4%)*	(9.7%)	(9.7%)	(3.7%)
All	25/135	61/135	21/135	38/135	23/135	15/135	22/122
	(18.5%)	(45.2%)*	(15.6%)	(28.2%)*	(17.0%)	(11.1%)*	(18.0%)

* Indicates significant difference at 5% compared to Actual Yes

Table III

Response patterns from Hypothetical and subsequent Actual purchasing questions

	Yes-Yes	Yes-No	No-Yes	No-No
200	11/36	9/36	1/36	15/36
	(30.56%)	(25%)	2.78%)	(41.67%)
250	7/33	9/33	0/33	17/33
	(21.21%)	(27.27%)	(0.00%)	(51.52%)
300	4/35	11/35	0/35	20/35
	(11.43%)	(31.43%)	(0.00%)	(57.14%)
400	2/31	8/31	0/31	21/31
	(6.45%)	(25.81%)	(0.00%)	(67.74%)
All	24/135	37/135	1/135	73/135
	(17.78%)	(27.41%)	(0.74%)	(54.07%)

Table IV

Proportion of Yes responses matching their Actual Yes responses (i.e. within respondent responses) across bid levels and groups

	Hypothetical Yes	Absolutely Yes	Certainty 8-10	Certainty 9-10	Certainty 10
200	55%	75%	67%	80%	100%
250	44%	100%	67%	100%	100%
300	27%	80%	50%	80%	75%
400	20%	67%	33%	67%	67%
All	39%	81%	58%	83%	87%

* Indicates significant difference at 5% compared to 'Actual Yes'

Table V
Non-parametric Turnbull WTP estimates

	Hypothetical – Actual sample						Actual Only sample
	Actual Yes	Hypothetical Yes	Absolutely Yes	Certainty 8-10	Certainty 9-10	Certainty 10	Actual Yes
WTP	89.44	189.0**	68.84	127.8	79.95	47.44*	83.14
	(16.92)	(19.52)	(15.45)	(18.66)	(16.42)	(14.13)	(17.41)

* significant difference at 10% when compared to Actual Yes
** significant difference at 5% when compared to Actual Yes