Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Health Economics 24 (2005) 515–529
JOURNAL OF HEALTH ECONOMICS
www.elsevier.com/locate/econbase

Sensitivity to scale in contingent valuation: the importance of the budget constraint

Richard D. Smith*

Health Economics Group, School of Medicine, Health Policy and Practice, University of East Anglia, Norwich NR4 7TJ, UK

Received 5 June 2003; received in revised form 5 April 2004; accepted 5 August 2004

Abstract

The possible insensitivity of willingness-to-pay (WTP) values to changes in scale continues to cause concern within the contingent valuation (CV) literature. Although several reasons for this phenomenon have been suggested, empirical evidence both supports and refutes insensitivity. This paper presents data that suggest that insensitivity may be significantly determined by the individual's budget constraint: the higher the proportion of income the expressed WTP represents, the greater the insensitivity of that WTP to changes in the scale of the good, *irrespective* of changes in underlying marginal utility. The methodological implication for CV studies in health care is outlined.

JEL classification: I10; C93; Q26

Keywords: Contingent valuation; Willingness-to-pay; Sensitivity to scale; Income

1. Introduction

Although adopted more extensively in environmental economics, valuing the benefits of health care interventions in monetary terms, using willingness-to-pay (WTP) in a contingent

* Tel.: +44 1603 593617; fax: +44 1603 593604.
E-mail address: richard.smith@uea.ac.uk.

0167-6296/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhealeco.2004.08.002

valuation (CV) survey, has increased significantly over the last decade (Diener et al., 1998; Klose, 1999; Olsen and Smith, 2001; Smith, 2003). However, a range of issues remain which require further research to refine the CV method, and thus improve the validity of WTP, including the (in)sensitivity of WTP to changes in the scale (amount) or scope (range) of the good being valued (Carson et al., 2001; Hanley et al., 2003).

Insensitivity means that WTP will not discriminate between different sizes/ranges of goods, and would also be said not to possess 'construct validity', as this result would not accord with economic theory of a positive correlation between WTP and benefit size/scope (Hammitt and Graham, 1999; Kartman et al., 1996). However, evidence of insensitivity within environmental economics is mixed (e.g. Arrow et al., 1993; Boyle et al., 1994; Carson, 1997; Diamond, 1996; Kahneman and Knetsch, 1992; Loomis and Larson, 1994; Smith and Osborne, 1995). The main conclusion drawn from this is that insensitivity indicates poor survey design, such as questions emphasizing the 'symbolic' nature of the good, problems in ensuring respondents comprehend probability, and/or small sample sizes generating a lack of statistical power (Bateman et al., 2002; Carson et al., 2001).[1] That is, the belief that the method is 'right', but is often applied wrongly.

However, there is no reason that environmental economics literature should necessarily apply in the health context (Olsen and Smith, 2001; Sheill and Gold, 2002; Smith, 2003). There are, for example, differences in the 'good' being valued – environmental goods are typically public goods (e.g. a species or area of wilderness) whereas health care goods are typically private (e.g. new treatment or screening test). Health economists therefore tend to be more skeptical that the CV method per se is 'right', notwithstanding whether it is applied well or not (Cookson, 2003). Yet, somewhat surprisingly, there has been little work on the sensitivity of WTP to scale/scope within the context of health care, especially to changes in health status per se (Hanley et al., 2003).

In two recent studies within the health care literature, sensitivity to scale[2] of WTP appeared to decrease as the size of the health benefit being valued increased (Smith, 2001; Yeung et al., 2003). In both cases, it was suggested that this may be due to an increasing 'relevance' of the budget constraint as the value of the good (relative to income) increases. That is, as the benefit increases, WTP for that benefit rises and consequently the budget constraint becomes an increasingly significant determinant of WTP. This would, of course, accord with the economic theory that underlies CV: that is, *Ceterus paribus*, more of a good is provided/consumed the more utility is gained, but that each successive increment in the good yields proportionally lower amounts of additional utility (the concept of diminishing marginal utility). Thus, in the indirect estimation of utility, where WTP is a proxy for utility, WTP should also increase at a decreasing rate for successive increases in health status (Flores and Carson, 1997). Depending upon the balance of these two factors (diminishing marginal

[1] Of course, the simple explanation is that people are 'genuinely' insensitive about the size of the good being valued. This would imply, for example, that respondents may be indifferent between an intervention expected to give them 10 extra years of life and one expected to give them 20 years. Although one has to be careful in ascribing 'rationality', this would appear to break even a minimum level of rationality. It is therefore taken as given that respondents *are* genuinely sensitive to variations in the scale/scope of goods being valued, and that there is thus some other explanation for why this may not be revealed by their WTP.

[2] Where 'scale' is defined as the level of health status (as a proxy for the quantity of 'health') 'possessed' by an individual, measured as quality-adjusted life years (QALYs).

utility and increasing 'impact' of the budget constraint), it may be that ability to pay, as approximated by income, is an important explanation of conflicting evidence concerning the observed (in)sensitivity of WTP to scale/scope. The mathematical expression of this issue is given in Appendix A.

In order to establish more directly the degree to which this issue is important, this paper tests the hypothesis that income is a key factor in determining sensitivity to scale *independent of changes in utility*. The study methodology and results are presented below, and the paper concludes with implications for the design and conduct of CV studies in health care and suggestions for further research.

2. Methods

The data presented in this paper are from a study in which WTP was assessed for a series of health state descriptions taken from a multi-attribute utility instrument developed in Australia: the Assessment of Quality of Life (AQoL) instrument (Hawthorne et al., 1999, 2000). The health states, valued using WTP, presented in this paper are a sub-set of those valued using the Time Trade-Off (TTO) method in the construction of the AQoL instrument (for a description of the TTO technique see Torrance, 1986). The focus of this paper is, however, upon the sensitivity of the WTP values to changes in the scale (amount) of 'utility' as approximated by the TTO values,[3] rather than a direct comparison between WTP and TTO values, which is reported elsewhere (Smith, 2001).

2.1. Health states valued

The AQoL comprises five dimensions of three items each, each with four levels of severity (from A to D, corresponding to best through worst health). This instrument is summarized in Table 1.

The health states (combinations of dimensions, items and levels from Table 1) presented in this paper are those that represent a range of absolute TTO values across the 0–1 scale (where 0 equates to being dead and 1 to being in good health). Ideally, a range of health states would be used that represent equivalent increments (e.g. each successive 0.1 increase) on this scale. However, the study was limited, as discussed below, to a selection of health states that had previously been valued directly using the TTO technique, and as such the increments represent *approximately* (rather than exactly) 0.1.

The health states, with their corresponding TTO values, are presented in Table 2. Note that the '1-TTO' figure provides the *gain* in health status that was valued using WTP (described below), and hence provides the comparator value for analysis (i.e. WTP is expected to increase as the 1-TTO figure increases). Note also that the 'extreme' values of moving from 0 or moving to 1 are not represented, as these were not asked.

[3] Of course the assumption here is that TTO values are a reasonable proxy for utility, and thus that a higher TTO value represents a greater 'quantity' of utility. If TTO values are not considered a reasonable proxy for utility in this sense, then this not only has serious implications for analysis presented in this paper but for the use of TTO and QALYs more generally.

Table 1
AQoL dimensions and items

Dimension	Item
1. Illness	1. Use of prescribed medicines 2. Reliance on medical aids 3. Receiving regular medical treatment
2. Independent living	4. Household tasks 5. Self-care 6. Mobility
3. Social relationships	7. Relationships with others 8. Social isolation 9. Family role
4. Physical senses	10. Seeing 11. Hearing 12. Communication
5. Psychological well-being	13. Sleep 14. Anxiety and depression 15. Pain

2.2. Sample

Respondents were drawn from a sample of 50 who had valued health states using the TTO technique approximately 4–8 weeks prior to the WTP survey described here. Although this was a pragmatic consideration, there are two advantages. First, the sample was recruited through random selection from Victorian electoral rolls within census collector divisions, stratifying for socio-economic status (Hawthorne et al., 1999, 2000). Thus, the results reported here represent a general population sample, rather than a specific patient, professional or other grouping. Second, as demographic and background data for these respondents already existed, this provided the opportunity to conduct a comprehensive, yet manageable, WTP interview, to help achieve a high response rate.

Table 2
Breakdown of scenarios used and associated TTO values

Question	Dimension	Item(s)	Severity[a]	TTO score (1-TTO score)
Q1	Illness	2	C	0.94 (0.06)
Q2	Independent living	5	C	0.87 (0.13)
Q3	Social relationships	7	D	0.78 (0.22)
Q4	Social relationships	7–9	D	0.72 (0.28)
Q5	Psychological well-being	13–15	D	0.60 (0.4)
Q6	Illness and independent living	1–6	D	0.49 (0.51)
Q7	Illness, independent living and social relationships	1–9	D	0.41 (0.59)
Q8	Illness, independent living, social relationships and physical senses	1–12	D	0.28 (0.72)

[a] Severity is classified according to a scale of A–D, where A is the least severe and D is the most severe.

2.3. Survey design and interview procedure

WTP values were derived using an 'open-ended' WTP question design, in which respondents were asked directly the maximum amount they would be willing to pay to move from a poor health state to good health.[4] The specific WTP question is provided in Appendix B. A prompt was used if respondents failed to provide a response within a period of approximately 60 s. This prompt was only used on three respondents, and then only for the first WTP question posed, and is thus an insignificant factor.

In administering the WTP questionnaire, respondents were first reminded of the previous TTO survey, and how they were asked to choose between different health states and express the time they would be prepared to give up to avoid a poor health state. They were told that this exercise would be the same, except that rather than giving up time to avoid a health state they would be required to give up money. The respondent was then presented with a series of choices between two health states. They were asked to assume they were in the 'poor health' state, and asked their *maximum* willingness-to-pay to secure a change from this state to the 'good health' state (i.e. a WTP to avoid being in the poor health state). Respondents were told that each health state (good or poor) was to last for a period of one year, after which time they would revert to 'good health'.

WTP was expressed as being paid out-of-pocket (OOP), either fortnightly (the traditional pay period for most people in Australia), monthly or annually as respondents felt comfortable. These values where then calculated as annual figures and repeated back to respondents for verification. Respondents were reminded to consider how much they could afford to pay when giving their response.[5]

No 'mechanism' by which the change in health state would be achieved was explicitly described. This was to try to remove any possible 'contamination' of the health state valuation according to the means by which that improvement in health state would be brought about (also reflecting the manner with which TTO values were obtained). If respondents asked how this change would be achieved, the interviewer explained that they should assume it would be a pill which was painless and involved no side-effects.

Prior to administration of the CV survey, the questionnaire was tested on a sample of 10 subjects taken from the AQoL sample (who were not interviewed during the main survey), who were subjected to additional questioning in a 'debriefing' session after the interview. This was undertaken to ensure that respondents understood the questions as posed and the implications of the valuations given, and to take account and address any ambiguities or problems that were felt to exist in the survey instrument. From these sessions, it was clear that respondents had little difficulty in understanding the questions, although there were some minor ambiguities in the questionnaire, which were addressed.

[4] The open-ended format was used as sample size constraints precluded the use of a dichotomous-choice format, and a simultaneous study by the author concerning different versions of the payment card format meant that it was unclear which of these versions would eventually be deemed most 'valid'. It is, however, noted that there remains debate concerning the relative merits of these different formats, although the open-ended and payment card formats may give approximate similar values (Smith, 2000; Frew et al., 2003).

[5] In contrast to many CV studies, in this survey respondents were *not* told that they would not actually have to pay this amount. As the purpose of the exercise was to establish the effect of income constraints, it was felt important that respondents perceive their constraint to be as 'real' as possible (Olsen and Smith, 2001).

A trained interviewer (who had also administered the TTO survey) conducted the interviews, at a time and place suitable for the respondent. Subjects were informed of the nature of the survey, both at initial contact via telephone to establish their willingness to participate, and also at the time of interview; where they also completed a consent form.

2.4. Statistical analyses

To assess the impact that income might have on sensitivity to scale, the degree to which WTP discriminated between each (approximately equal) change in TTO score through the range from 0 to 1 was tested. There is no precedent for a 'meaningful' difference in TTO score, but here it is considered to be approximately 0.1. Thus, the hypothesis is that, for example, the difference in WTP values corresponding to a change from 0.1 to 0.2 on the TTO scale may be significant, and thus 'pass' a sensitivity to scale test, but that the difference in WTP values for an equivalent change from 0.6 to 0.7 may not be significant, thus 'failing' a sensitivity to scale test. As the 'objective' measure of utility increment (TTO) is the same,[6] this might then be assumed to represent an impact of the budget constraint (WTP becomes less sensitive the higher the quantity (scale) of the good (health state improvement) being purchased). As the data were negatively skewed, a non-parametric statistical analysis, the Wilcoxon paired comparison (signed rank) test, was used to assess these differences (Siegal and Castellan, 1988).

As a second test, multiple linear regression was used to test for the impact of income, controlling for several other factors, upon the WTP values for each question. If income is a key factor in determining WTP, as suggested above, then one would expect income to be a significant explanatory variable. The model was specified to be parsimonious, yet encompass the main variables that might be expected to influence WTP, including, in addition to income, age, gender, whether the respondent was cohabiting and whether they had dependent children (i.e. were caring for significant others). Self-reported current health status was not included in these analyses as it was found to be closely correlated to income. Given the skewed nature of the WTP data, as indicated, a log-linear (semi-log) regression model was used, where the dependent variable, WTP, is log transformed (Gujarati, 1988). A further analysis pooled the WTP data and incorporated a variable to reflect the quantity of health benefit and interaction between this and income. All analyses were conducted using Stata for Windows, version 8.0.

3. Results

3.1. Sample size, response rate and demographic characteristics

Between August 1996 and June 1997, 47 interviews were completed of the 50 individuals approached, a recruitment response rate of 94%. Respondents were on average aged 43 years (±1.9), their average income was AU$46,170 (±AU$4013), 47% were male, 89%

[6] To enable the calculation of quality-adjusted life years, a 0.1 increment is required to be equivalent in value regardless of where on the 0–1 scale it occurs (Drummond et al., 1997).

Table 3
WTP values

Question	Mean	%WTP[a]	Median	S.D.	Skew	Minimum	Maximum	S.E.	95% CI
Q1	1020	1.8	500	1435	1.66	0	5200	216	584æ1457
Q2	1498	3.4	1040	1531	2.07	130	7800	231	1032æ1963
Q3	4845	9.4	2800	5106	1.99	520	25000	770	3292æ6397
Q4	7607	14.5	5200	7443	2.86	520	40000	1122	4869æ10334
Q5	9417	18.8	7800	9023	2.42	1200	50000	1360	6674æ12160
Q6	9535	19.4	7800	9560	2.75	1200	55000	1441	6629æ12442
Q7	9602	19.7	7800	9527	2.77	1200	55000	1436	6706æ12499
Q8	9773	19.8	7800	10110	3.05	1200	60000	1524	6699æ12846

[a] Mean WTP expressed as a proportion of income.

rated their current health as ≥'average' and 83% were married or in de facto relationships. Interviews took an average of 65 min (±2.8) to complete.

3.2. WTP values

Table 3 provides a comprehensive breakdown of WTP response by question. Although, as with most WTP surveys, there is a skewed distribution of values, mean and median values remain relatively close (note that 15 respondents gave a 0 WTP value for Q1, none of which was determined to be a 'protest' bid, and there were no other zero bids).

Table 4 presents the result of the Wilcoxon test for the association between WTP and TTO for pairs questions. It provides the raw difference in TTO (an approximate 0.1 difference in TTO), the raw difference in mean WTP (and WTP as a proportion of income) associated with it and a test of significance of the null hypothesis that this difference in mean WTP was not statistically different (i.e. that the WTP was insensitive to this specific change in scale of the good (health state) being valued).

For each movement of 'utility' along the 0–1 scale of approximately 0.1, the equivalent change in WTP is significantly different for absolute TTO values of less than 0.6 (Q5), but is insignificantly different for values over this. Thus, if the WTP values had only been collected for the 'better' health states (Q1–Q4) then these values would have been judged to be sensitive to scale (WTP would have discriminated between successive increases in the

Table 4
Sensitivity to scale: significance of association between WTP values corresponding to different TTO values

TTO questions	Change in mean TTO	Change in mean WTP	Change in %WTP	Significance level (*P*-value)
Q1 vs. Q2	0.07	478	1.6	<0.001
Q2 vs. Q3	0.09	3347	6.0	<0.001
Q3 vs. Q4	0.06	2762	5.1	<0.001
Q4 vs. Q5	0.12	1810	4.3	<0.001
Q5 vs. Q6	0.11	118	0.6	0.157
Q6 vs. Q7	0.08	67	0.3	0.056
Q7 vs. Q8	0.12	171	0.1	0.083

Table 5
Predicted WTP values on the basis of linearity

Question	Actual WTP	Predicted WTP from Q5[a]	Predicted WTP from Q4[b]
Q1	1020	–	–
Q2	1498	–	–
Q3	4845	–	–
Q4	7607	–	–
Q5	9417	–	11934
Q6	9535	11677	18736
Q7	9602	14480	29415
Q8	9773	17955	46182

[a] WTP values here are increased in line with the proportion increase seen from Q4 to Q5 (a rate of approximately 24%).
[b] WTP values here are increased in line with the proportion increase seen from Q3 to Q4 (a rate of approximately 57%).

utility derived from the health states). However, if the WTP values had only been collected for the 'worse' health states (Q7–Q8), then these values would have been judged to be insensitive to scale (WTP invariant to changes in utility).

Table 5 thus attempts to 'predict' what WTP 'ought' to be if linearity held with respect to WTP as well as to TTO, using the rate of increase from the last significant change in WTP (that is, from the range that were judged to be sensitive to scale).

Based on the rate of increase suggested by the additional WTP to 'avoid' Q5 than Q4 (a rate of 24%), WTP is predicted to increase to just under $18,000 – approximately twice the actual WTP given for Q8. If this predicted WTP is based on the additional WTP suggested by the move from Q3 to Q4, WTP would be predicted to increase to approximately the same level as the average income of the sample – just over $46,000.

Table 6
Variables used in regression analyses

Variable	Description	Type of variable
LnWTP*x*	Log of maximum annual willingness-to-pay for question *x* (where *x* = 1–8)	Continuous
LnWTPlp	Log of maximum annual willingness-to-pay for the pooled sample	Continuous
LnWTPl	Log of maximum annual willingness-to-pay for changes in TTO at a low level (Q1–Q4)	Continuous
LnWTPh	Log of maximum annual willingness-to-pay for changes in TTO at a high level (Q5–Q8)	Continuous
Age	Age of respondent	Continuous
Gender	Female = 1	Dummy variable
Income	Annual income of respondent	Continuous
Cohabit	Cohabitation status, where 1 = cohabiting	Dummy variable
Kids	Dependent children = 1	Dummy variable
QHS	Quantity of health status (1-TTO value)	Continuous
Inc × QHS	Interaction term for income and QHS	Continuous

Table 7
Regression results (S.E.)

Variable	LnWTP1	LnWTP2	LnWTP3	LnWTP4	LnWTP5	LnWTP6	LnWTP7	LnWTP8	LnWTPp	LnWTPl	LnWTPh
Cons	+2.270* (0.08)	+2.602* (0.042)	+6.654* (0.005)	+1.752* (0.053)	+2.833* (0.022)	+3.153* (0.016)	+3.153* (0.015)	+3.503* (0.011)	+3.37* (0.028)	+2.835* (0.010)	+2.069* (0.052)
Age	+0.002 (0.002)	−0.067 (0.000)	+0.017 (0.000)	+0.001 (0.001)	−0.000 (0.000)	−0.000 (0.000)	−0.000 (0.000)	−0.000 (0.000)	−0.001 (0.006)	−0.001 (0.013)	−0.001 (0.000)
Gender	−0.015 (0.042)	+0.003 (0.259)	+0.000 (0.001)	−0.016 (0.032)	+0.008 (0.013)	+0.000 (0.010)	+0.005 (0.009)	+0.003 (0.007)	+0.048 (0.194)	+0.044 (0.393)	+0.006 (0.013)
Income	+0.032* (0.078)	+0.016* (0.004)	+0.039* (0.053)	+0.040* (0.005)	+0.018* (0.021)	+0.019* (0.015)	+0.017* (0.014)	+0.082* (0.010)	+0.019* (0.059)	+0.061* (0.011)	+0.020* (0.008)
Cohabit	+0.112** (0.067)	+0.042 (0.035)	+0.001** (0.004)	+0.037 (0.043)	+0.007 (0.018)	+0.005 (0.013)	+0.008 (0.008)	+0.007 (0.009)	+0.028 (0.062)	+0.054** (0.034)	+0.007 (0.017)
Kids	−0.035 (0.037)	−0.017 (0.024)	+0.000 (0.000)	−0.006 (0.029)	−0.013 (0.012)	−0.009 (0.009)	−0.009 (0.009)	−0.007 (0.006)	−0.023 (0.076)	−0.027 (0.355)	−0.009 (0.012)
QHS	–	–	–	–	–	–	–	–	+0.084 (0.729)	+0.966** (0.088)	+0.005 (0.974)
Inc × QHS	–	–	–	–	–	–	–	–	+0.029 (0.100)	+0.025 (0.081)	+0.001 (0.005)
Adj R^2	0.37	0.24	0.44	0.51	0.52	0.52	0.52	0.53	0.32	0.48	0.40

* Significant at 5%.
** Significant at 10%.

3.3. Regression

The regression model was specified using the variables as defined in Table 6.

This model was run for each WTP question (Q1–Q8), with a further analysis based on pooled data. Results are presented in Table 7.

Income is positively associated with the (log) WTP value for each individual Q1–Q8. No other variables are significant across all questions, with the exception of cohabitation status for Q1 and Q3. All variables have a positive effect on WTP, with the exception of whether the respondent has children. These seem an intuitive direction of relationship. For example, cohabiting may reflect a wider access to funds than personal income provides (only personal income was asked for in the interview), whereas having children may be a drain on income. If anything, it is surprising that these factors were not significant across more questions. For Q4 and above, the model is explaining just over 50% of the variance in WTP, which is surprisingly high.

These broad results for individual questions also hold for the pooled regression – income is the only significant explanatory factor. On this basis, WTP appears to be insensitive to scale as QHS is not significant. However, split sample results suggest that insensitivity only applies where the change in health status (QHS) is at a high level (representing Q5–Q8) – LnWTPh shows income but *not* QHS as significant variables – and not where this change is at a low level (representing Q1–Q4) – LnWTPl shows *both* income *and* QHS as significant variables. These results reinforce the central hypothesis advanced in this paper for the income effect on sensitivity. (Note that the interaction term in all pooled analyses is not significant, indicating that the simultaneous presence of income and QHS does not attenuate or reinforce their individual effects.)

4. Conclusion

There are, of course, several caveats to be placed on these results. For example, the pattern of values demonstrated here may be a peculiar result of the sample size or characteristics, the specific methodology used or the 'good' being valued (change in health state).[7] However, although the results of this study are tentative, rather than definitive, they do indicate that it may be *possible* for studies to show sensitivity or insensitivity to scale for the same *marginal* changes in benefit simply because they differ where they are on the *absolute* scale over which changes in WTP are being valued.

If replicable, this result has an important methodological implication for CV studies in health care. In general, compared with environmental economists, health economists work with small samples of patients, or populations at risk (Olsen and Smith, 2001). The consequent lack of statistical power, coupled with the properties of diminishing marginal WTP values, could lead one to reject WTP estimates based on a false negative inference with respect to insensitivity to scale. This places the emphasis firmly on the recruitment of a

[7] Of course, these results may be consistent with a linear translation *between* TTO and WTP, but non-linearity *within* TTO. However, as indicated earlier, the assumption in this paper of linearity in changes in TTO is based on that required to enable the calculation of QALYs.

sample size sufficient to be able to detect significant changes in WTP across changes in the size of benefit (insofar as one is testing for such significance of course). Where WTP (as a proportion of income) is expected to be high, this will necessitate a far larger sample size that is usually recruited in CV studies in health care, and one that is sufficiently representative of heterogeneity in income.

Clearly, more research is required, not only to assess the replicability of these results, but also to assess at what level of WTP, as a proportion of income, sensitivity to scale is likely to be compromised. From Table 3, it would appear from this study that this 'threshold' income level may be quite low – although at a diminishing rate, WTP as a proportion of income increases quite rapidly across the 'low' range of TTO values, reaching approximately 19% at a 1-TTO value of 0.4. However, it then increases by only 1%, to approximately 20%, during the remaining range to a 1-TTO of 0.72. As indicated above, this could be a function of study design (for example, different elicitation formats may produce different upper bounds (Whynes et al., 2004)), but 20% of income as an 'upper bound' for WTP does seem low. On the face of it, this might reflect a general level of disposable, or discretionary, income once fixed costs of living, such as mortgages and food, have been accounted for. However, this assumes that health is a 'luxury' good; something which is bought *after* these other costs of living are paid for. Although possible, this may be counter-intuitive – that health, as a 'fundamental commodity' would be 'bought' first – and a different explanation required. It could also be that the solution is determined simultaneously – respondents make adjustments for their *expected* commitments on the basis of the level of health they are *expecting* once they have bought it.

Sensitivity to scale may be an area where CV studies in health and the environment differ, as the former value predominantly use value and the latter predominantly non-use value. For example, use value may generate higher WTP values (as a proportion of income) than non-use value, suggesting that the income constraint may be a more acute effect in health than environmental CV. If so, then one may expect insensitivity to scale to be a greater issue in health than environmental valuation, and CV studies in health will, *C. paribus*, require larger sample sizes that CV studies in environmental economics. The opposite could also be the case, but either way the proposition outlined in the introduction, that one cannot *necessarily* take and apply findings in the environmental literature to CV studies in health, still holds and further consideration of these methodological issues is required in health if the potential of CV is to be realized.

Acknowledgements

The research described in this paper was conducted whilst the author was working at the Health Economics Unit, Faculty of Business and Economics, Monash University, and was funded by a Faculty Research Grant. The author would like to thank Jeff Richardson and Graeme Hawthorne for providing comments on the ideas contained in this paper and allowing the use of AQoL TTO data, Helen McNeil for conducting interviews, David Whynes, Paula Lorgelly and an anonymous referee for comments on an earlier draft and, of course, to all those who completed the interviews. Needless to say, the usual caveat applies.

Appendix A. Mathematical expression of sensitivity issue

Following Rollins and Lyke (1988), the sensitivity issue may be expressed as

$$\mathrm{WTP}(\Delta h) = e(p, h, u) - e(p, h + \Delta h, u) \tag{1}$$

$$\frac{\partial \mathrm{WTP}(\Delta h)}{\partial h} = \frac{-\partial e(p, h + \Delta h, u)}{\partial h} \tag{2}$$

where $e(p, h, u)$ denotes the expenditure function that maximizes utility (u), given an endowment of health (h) and the price of all other goods (p). This changes as the endowment of health changes (Δh). The right hand side of Eq. (2) is the utility-constant marginal value, such that

$$\frac{\partial \mathrm{WTP}(\Delta h)}{\partial h} = \frac{-\partial e(p, h + \Delta h, u)}{\partial h} = v(p, h + \Delta h, u) \tag{3}$$

where the marginal value (v) is positive but diminishing.[8] That is, WTP will increase at a decreasing rate for successive increases in health status (h). The goal of a sensitivity test is to establish a positive value of $\frac{\Delta \mathrm{WTP}}{\Delta h}$. The representation of the integral implies that

$$\frac{\partial \mathrm{WTP}}{\partial h} = p^{\mathrm{h}}(p, h, u) \tag{4}$$

(where p^{h} is the 'virtual price' of health), which is a positive number (for a positive Δh, ΔWTP is positive). However, the 'validity' of diminishing WTP depends upon it reflecting an underlying diminishing *utility*, where an observed diminishing marginal WTP: benefit size ratio could also reflect an increasing 'impact' of the budget constraint (Donaldson, 1999). Depending upon the balance of these two factors, it may be that ability to pay, as approximated by income, is an important explanation of conflicting evidence concerning the observed (in)sensitivity of WTP to scale/scope. That is, the presence of a budget constraint prohibits some respondents from exhibiting a positive $\frac{\Delta \mathrm{WTP}}{\Delta h}$. Under a virtual price representation, the outcome of interest is therefore whether

$$\frac{\partial^2 \mathrm{WTP}}{\partial h \partial y} = \frac{\partial p^{\mathrm{h}}(p, h, u)}{\partial y} \geq 0. \tag{5}$$

Thus, for each individual the size of $\frac{\Delta \mathrm{WTP}}{\Delta h}$ is larger when evaluated at a lower level of h than at a higher level because of the constraint placed on WTP.

Appendix B. WTP question

The questions will all follow the same format. You will be presented with a pair of health states. In each case I will ask *you to assume that you are in health state B* [*Good Health*], and

[8] The second derivative of WTP will be equal to the derivative of the marginal value (v) with respect to Δh, which in the case of a single good is always negative for convex preferences (Madden, 1991).

will ask you the maximum amount you would be willing-to-pay *to avoid going immediately into state A*, as presented. Please think of this willingness-to-pay as being directly paid out-of-pocket. This payment will be ongoing and have to be made regularly. If it stops you will go into state A immediately. You can think of this regular willingness-to-pay as either a weekly, fortnightly, monthly or yearly figure.

Also, for the purpose of this interview, please imagine that there will be no help with paying for this treatment from Medicare. If you want to avoid going into health state A you will have to pay for treatment yourself. We want to know how much you would be willing to pay.

When giving your response, that is, what you are willing to pay, please bear in mind what you are able to afford. Given your own circumstances (assets and income), bear in mind that if you are willing to pay an amount to secure this improvement you will have less to spend on other things, such as entertainment, your house, car, holidays and so on.

> If you are not willing to pay anything to change health states, then please say zero.
>
> As always there are not right or wrong answers. We are only interested in what you think.
>
> [*Interviewer*: Get each set of health states out one by one, and offer the choice to respondents. Record the maximum willingness to pay for each one in the space provided and also record whether or not the prompt was used.].

If the respondent is unable to give an amount then this prompt may be used and the 'prompt used' circled.

PROMPT

Would it help if I suggested an amount to give you a start?

Would you be willing to pay $X [vary X across $25, $50, $75, $100]

If no, then read out decrements in $10 amounts steadily and quickly.

If yes, read out increments in $10 amounts steadily and quickly.

When respondent stops, ask

So you would be WTP $X pw to be in HS B rather than HS A?

Would you pay exactly X or slightly more or less. How much exactly?

WRITE IN SUM $. . .PROMPT USED

If any response is a zero, ask the respondent to clarify if they are simply not willing to pay anything, that they do not actually value the change in health state (i.e. are indifferent between states A and B), or there is another reason – such as income constraint, ethical objections to such valuation, feel that this payment may become policy, etc.

References

Arrow, K., Solow, R., Portney, P., Leamer, E., Radner, R., Schuman, H., 1993. Report of the NOAA panel on contingent valuation. Federal Register 58, 4601–4614.

Bateman, I.J., Carson, R.T., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., Loomes, G., Mourato, S., Ozdemiroglu, E., Pearce, D.W., Sugden, R., Swanson, R., 2002. Economic Valuation With Stated Preference Techniques: A Manual. Edward Elgar, Cheltenham, pp. 296–342 (Chapter 8).

Boyle, K.J., Desvousges, W.H., Johnson, F.R., Dunford, R.W., Hudson, S.P., 1994. An investigation of part-whole biases in contingent valuation studies. Journal of Environmental Economics and Management 27, 64–83.

Carson, R.T., 1997. Contingent valuation surveys and tests of insensitivity to scope. In: Kopp, R.J., Pommerehne, W.W., Schwarz, N. (Eds.), Determining the Value of Non-marketed Goods: Economics, Psychology and Policy Relevant Aspects of the Contingent Valuation Method. Kluwer Academic, Dorsrecht, The Netherlands.

Carson, R.T., Flores, N.E., Meade, N.F., 2001. Contingent valuation: controversies and evidence. Environmental and Resource Economics 19, 173–210.

Cookson, R., 2003. Willingness to pay methods in health care: a skeptical view. Health Economics 12, 891–894.

Diamond, P.A., 1996. Testing the internal consistency of contingent valuation surveys. Journal of Environmental Economics and Management 30, 337–347.

Diener, A., O'Brien, B., Gafni, A., 1998. Health care contingent valuation studies: a review and classification of the literature. Health Economics 7, 313–326.

Donaldson, C., 1999. Valuing the benefits of publicly-provided health care: does 'ability to pay' preclude the use of 'willingness to pay'? Social Science and Medicine 49, 551–563.

Drummond, M.F., O'Brien, B., Stoddart, G., Torrance, G., 1997. Methods for the Economic Evaluation of Health Care Programmes, second ed. Oxford University Press, New York.

Flores, N.E., Carson, R.T., 1997. The relationship between the income elasticities of demand and willingness to pay. Journal of Environmental Economics and Management 33, 287–295.

Frew, E., Whynes, D., Wolstenholme, J., 2003. Eliciting willingness to pay: comparing closed-ended with open-ended and payment scale formats. Medical Decision Making 23, 150–159.

Gujarati, D., 1988. Basic Econometrics. McGraw-Hill, Singapore (Chapter 6).

Hammitt, J.K., Graham, J.D., 1999. Willingness to pay for health protection: inadequate sensitivity to probability? Journal of Risk and Uncertainty 8, 33–62.

Hanley, N., Ryan, M., Wright, R., 2003. Estimating the monetary value of health care: lessons from environmental economics. Health Economics 12, 3–16.

Hawthorne, G., Richardson, J., Osborne, R., 1999. The Assessment of Quality of Life (AQoL) instrument: a psychometric measure of health related quality of life. Quality of Life Research 8, 209–224.

Hawthorne, G., Richardson, J., Day, N., Osborne, R., McNeil, H., 2000. Construction and utility scaling of the assessment of Quality of Life (AQoL) instrument. Working Paper 101, Centre for Health Program Evaluation, Monash University, Melbourne, Australia.

Kahneman, D.J., Knetsch, J., 1992. Valuing public goods: the purchase of moral satisfaction. Journal of Environmental Economics and Management 22, 57–70.

Kartman, B., Stalhammar, N.O., Johannesson, M., 1996. Valuation of health changes with the contingent valuation method: a test of scope and question order effects. Health Economics 5, 531–541.

Klose, T., 1999. The contingent valuation method in health care. Health Policy 47, 97–123.

Loomis, J., Larson, D., 1994. Total economic values of increasing gray whale populations: results from a contingent valuation survey of visitors and households. Marine Resource Economics 9, 275–286.

Madden, P., 1991. A generalization of Hicksian q substitutes and complements with application to demand rationing. Econometrica 59, 1497–1508.

Olsen, J.A., Smith, R.D., 2001. Theory versus practice: a review of 'willingness-to-pay' in health and health care. Health Economics 10, 39–52.
Rollins, K., Lyke, A., 1988. The case for diminishing marginal existence values. Journal of Environmental Economics and Management 36, 324–344.
Sheill, A., Gold, L., 2002. Contingent valuation in health care and the persistence of embedding effects without the warm glow. Journal of Economic Psychology 23, 251–262.
Siegal, S., Castellan, N.J., 1988. Nonparametric Statistics for the Behavoural Sciences, second ed. McGraw Hill, Singapore.
Smith, R.D., 2000. The discrete choice willingness-to-pay question format in Health Economics: should we adopt environmental guidelines? Medical Decision Making 20, 194–206.
Smith, R.D., 2001. The relative sensitivity of willingness-to-pay and time-trade-off to changes in health status: an empirical investigation. Health Economics 10, 487–497.
Smith, R.D., 2003. Construction of the contingent valuation market in health care: a critical assessment. Health Economics 12, 609–628.
Smith, V.K., Osborne, L., 1995. Do contingent valuation estimates pass a "scope" test? A meta-analysis. Journal of Environmental Economics and Management 31, 287–301.
Torrance, G.W., 1986. Measurement of health state utilities for economic appraisal: a review. Journal of Health Economics 5, 1–30.
Whynes, D.K., Wolstenholme, J.L., Frew, E.J., 2004. Evidence of range bias in contingent valuation payment scales. Health Economics 13, 183–190.
Yeung, R., Smith, R.D., McGhee, S.M., 2003. Willingness to pay and size of health benefit: an integrated model to test for 'sensitivity to scale'. Health Economics Letters 12, 791–796.