DEPARTMENT OF ECONOMICS

SCHOOL OF SOCIAL SCIENCES

A theoretical framework for TTO valuations and a taxonomy of TTO approaches: results from a pilot study

Ken Buckingham, Nancy Devlin* & Maggie Tabberer
City Health Economics Centre, Economics Department, City University.

Department of Economics
Discussion Paper series
No. 04/07

Correspondence to:
Nancy Devlin
City Health Economics Centre
Economics Department
City University.
Northampton Square
London EC1 V OHB.
e-mail: n.j.devlin@city.ac.uk
phone: 0044 (0)20 7040 8518
fax: 0044 (0)20 7040 8580

A theoretical framework for TTO valuations and a taxonomy of TTO approaches: results from a pilot study

1 Introduction

The economic evaluation of health care therapies has come to be dominated by Cost Utility Analysis (CUA), usually performed estimating the incremental cost per Quality Adjusted Life Year (QALY) gained of each option. The use of QALYs requires the valuation of each relevant health state on a scale anchored at 1 (full health) and 0 (dead); commonly used techniques for eliciting these valuations include Standard Gamble (SG), Time Trade-off (TTO) and Visual Analogue Scale (VAS).

While SG is rooted in von Neumann Morgenstern utility theory (von Neumann and Morgenstern 1944, Drummond et al 1997), the theoretical underpinnings of other valuation methodologies are less clear. Direct valuation of health states using VAS, for example, is widely considered to lack a theoretical basis (Johannesson et al 1996; for a contrasting view, see Parkin and Devlin 2004). The TTO method arose not from any particular theory of utility or utility measurement, but rather from a pragmatic desire to generate valuations with similar empirical properties to those generated from SG, while offering a more feasible alternative to it. By convention, TTO valuations for a given health state H_i are elicited by asking subjects to consider t time in that state; and then establishing the (lower) number of years, x, in full health, they would be consider equivalent; the value for H_i is given by x/t. While TTO valuations do not conform to the utility-under-uncertainty requirements of expected utility theory, and are therefore considered to be valuations rather than utilities *per se* (Drummond et al 1997), they have nevertheless come to dominate the valuation sets routinely used in CUA in the UK (for example, by NICE), the US and elsewhere to inform health sector priority setting and resource allocation decisions.

Given their widespread application in economic evaluation, the empirical properties of TTO valuation data have been subject to extensive investigation but, arguably, there have been few attempts to locate the approach within any particular body of theory. The aims of this paper are to consider the specific approach conventionally used to elicit TTO values within a body of utility theory and, by extension, to consider what other approaches to the TTO task are also suggested by that theory. We report results from a pilot study conducted to investigate the feasibility of the alternative approaches and the characteristics of the valuation data generated by them.

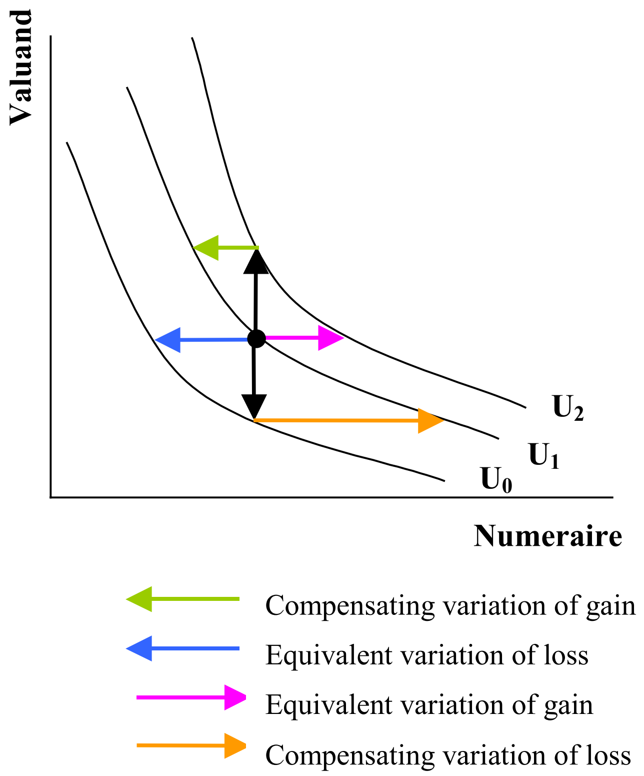
1.1. Hicks utility theory

Hicks (1946) describes compensated demand curves and the possibility of valuing welfare changing events (such as changes in prices or endowments of economic goods) by determining the compensating change in income that would return the consumer to their original utility level. Such valuations are referred to as Compensating Variations (CV). Hicks also describes an alternative method in which the value of a particular welfare change might be assessed as an equivalent welfare effect achieved instead by a change in income. In this way the focus is moved to the level of utility experienced *after* the change in events. These valuations are referred to as Equivalent Variations (EV). The distinctive properties of EV compared to CV as monetary measures of the utility affords each specific advantages in particular contexts (Johansson 1991); both suffer from problems in aggregation for the purpose of hypothetical compensation tests (Johansson 1991; Mishan 1971). Neither measure can be regarded as strictly superior to the other (Ng 1979).

Both compensating and equivalent variations can be calculated for welfare gains and welfare losses. Using the postscripts 'G' and 'L' to indicate respectively 'gains' and 'losses' in health, this implies four valuation possibilities; CVG, CVL, EVG and EVL, as shown in Figure 1. We use the term 'valuand' generically to describe the 'thing being valued', whether that 'thing' be goods, services or intangibles. Both the CV and the EV derive from a conceptualisation of utility space in which fixed levels of utility are associated with combinations of commodities. The rate at which changes occur and the sequence of such changes are assumed to be irrelevant to valuations.

For a variety of reasons, health economists have sought alternatives to income (or wealth) as the numeraire applied to the valuation of changes in health. Thus, Cost Benefit Analysis, which requires monetary assessments of 'benefit', is generally eschewed in favour of CUA. The values applied in the estimation of QALYs have used both reductions in survival probabilities (SG) and reductions in life expectancies (TTO) as the numeraire with which changes in health-related quality of life (hereafter referred to as 'health') are evaluated.

Figure 1. Valuations of losses and gains by compensating and equivalent variations



The conventional TTO procedure described above can be considered as a CVG i.e., the reduction in life expectancy (the numeraire) required to return the individual to their pre-existing utility following a positive change (a 'gain') in health (the valuand). We have already noted that there are no prior theoretical grounds for favouring CVG or CVL over their EV counterparts; consequently all four valuation methods provide theoretically possible ways of valuing a given change in health, as described below.

The TTO as a Compensating Variation for a Health Gain (CVG): The conventional approach to TTO (Torrance, 1976).

The TTO as a Compensating Variation for a Loss (CVL): We could value the same change in health, with the same initial endowment of life expectancy, but as a *reduction* from perfect health to a given state of ill health by establishing the compensating *increase* in longevity required to return the individual to their original level of utility. CVL and CVG cover different areas in utility space and there is no reason to suppose that the valuations would be equal (see Figures 2a and 2b).

Figure 2a.
CV & EV applied to the same gain in health

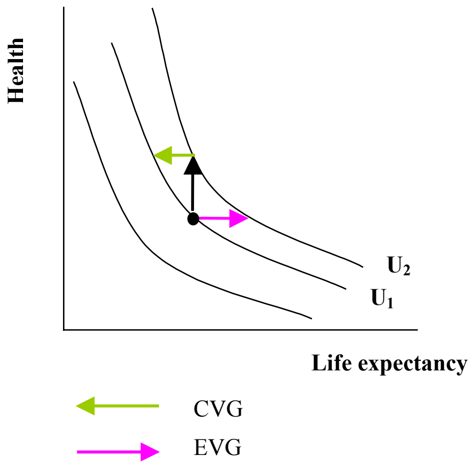
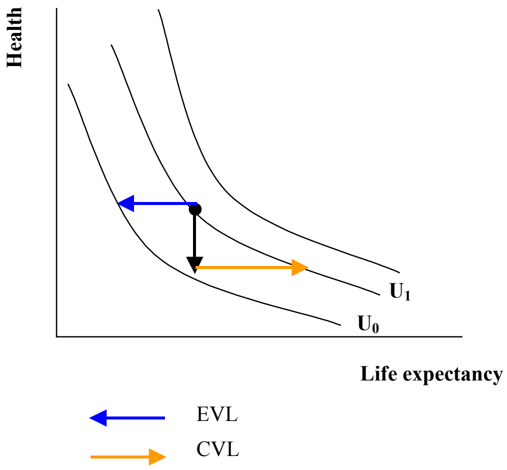


Figure 2b
CV &EV applied to the same loss in health



The TTO as an Equivalent Variation for a Gain (EVG): The possibility of using an equivalent variation to value an improvement in health was suggested by Buckingham (1995). An equivalent variation for a gain in health would attempt to determine the *increase* in life expectancy that would take the respondent to the same (higher) utility level as the *increase* in health. Careful inspection of Figures 2a and 2b will show that the utility space that such a question spans is identical to that spanned by the compensating variation for the same change in health viewed as a health loss (CVL). In the absence of any framing effects (e.g., resulting from the valuation of a loss as opposed to an identical gain – we discuss a number of such effects in Section 1.2), there is no reason *a priori* to expect that the valuations yielded by EVG and CVL would be different.

The TTO as an Equivalent Variation for a Loss (EVL): An equivalent variation valuation of a reduction in health over the same range is also possible and spans the same utility space as the conventional compensating variation valuation of a health gain (CVG). Thus, again in the absence of framing effects, we would expect these measures to produce the same valuation.

Note that under some TTO protocols a starting point combination of life expectancy and health state is deliberately not specified to avoid any 'endowment effect'. In the absence of such a starting point the classification of technique as either CVG or EVL is indeterminate, since the question of whether we are considering alternative ways of obtaining a different level of utility i.e., EV, or returning to a prior level of utility ie CV is avoided. Nevertheless, we would argue that the approach is still fundamentally Hicksian, since the aim is to identify points of indifference to achieve valuations.

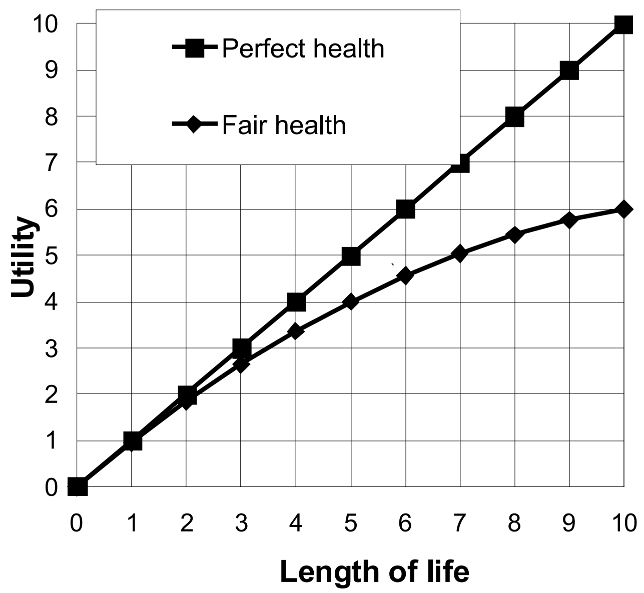
1.2 The calculation and interpretation of a health state index

The application of Hicksian welfare theory to health is somewhat different from its more familiar uses, most obviously because tradeoffs are sought in terms of life years rather than money. More fundamentally, the ultimate aim is to produce health state indices rather than simply valuing welfare changes. In Figure 3 (an alternative representation of the information conveyed using indifference curves) we see the change from fair health to perfect health is compensated by a reduction in life expectancy from 10 years to 6 years. The compensating variation would thus be 4 years. Two additional steps are required to convert this compensating variation into an index. (Analogous procedures are necessary for equivalent variations.)

Firstly we must assume that utility is proportional to time spent in perfect health. Arbitrarily we define our measure of utility, such that 1 year in perfect health is equal to 1 'util'. With proportionality x years in perfect health is equivalent to x utils. In Figure 3, we see that 10 years in fair health has the same utility as 6 years in perfect health, and hence the utility of 10 years in fair health is 6 utils.

One final step is required to arrive at an index. A requirement of the index is that it can be multiplied by time to construct QALYs where QALYs are used as a proxy measure of utility. The only possible interpretation of something that we multiply by time to produce a measure of utility is that it is a flow of utility per period of time. To convert our proxy for the absolute value of utility of fair health for 10 years, into an average flow of utility over those years, we must divide it by the time taken to achieve the utility. Thus in our illustrated example, the index is (6 utils)/(10 years) or 0.6 utils per year.

Figure 3. The application of proportionality in constructing a health index



Note that if this value is to be applicable to different time periods, we must make a further proportionality assumption. We must now assume that utility is proportional to the time spent in the state of health that has been valued for all life expectancies. Unless this were the case, the average flow for any given period would not necessarily be applicable to other periods. In Figure 3, we show a non- linear relationship for fair health such that utility is not proportional to the time spent in that state. If we were to try to estimate the utility of 5 years in fair health by multiplying those 5

years by our index of 0.6, we would estimate the value of 5 years in fair health as 3 utils. In Figure 3 the utility shown is actually 4 utils, i.e. the QALY value (product of the index multiplied by time) gives an incorrect estimate of utility.

Thus, to generate a TTO index that is applicable to all time periods, we need to employ the assumption of proportionality twice. Firstly, in assuming proportionality at perfect health, and secondly assuming proportionality at less than perfect health i.e., corresponding to the two lines shown n Figure 3.

Note that the Standard Gamble also needs to employ two proportionality assumptions to produce an index, but the first of those assumptions is proportionality of utility to probability rather than time. This is used to value the health state. The assumption of proportionality of utility to time in the health state being valued is subsequently used to convert the value of that health state into an index. These two steps may be masked by the scaling assumptions used to construct the Standard Gamble.

1.3 Recent TTO research in the context of this theoretical framework

Alternative formulations of the TTO task have received increasing attention in the literature. In this section we review recent empirical work on TTO with the purpose of (a) placing these findings in the context our theoretical model and (b) deriving hypotheses relevant to the characteristics of individuals' utility functions between health and length of life that might be explored in a pilot study of the four approaches we describe above.

Empirical studies have usually revealed differences in the TTO valuations given to health states when alternative formats of the valuation task are used. The most described effect is that of time preference. For example, Dolan and Gudex (1995) attempted to separate the effects of time preference from that of duration and concluded that TTO is not suitable for the valuation of short duration health states where the life expectancy presented is unrealistic to the respondent.

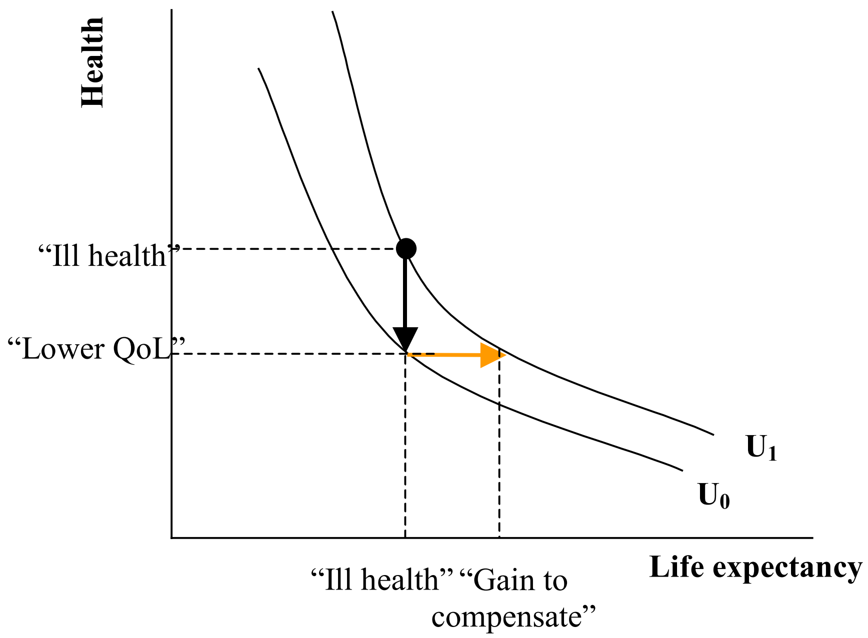
The conventional (CVG) TTO valuation assumes linear increases in utility for time spent in perfect health. SG does not assume this linearity and this has been used to explain why elicited TTO valuations are lower than SG valuations for the same state. Consistent valuation over the full range of health states assumes a constant proportional trade off. Dolan and Stalmeier (2003) concluded that this may not always be the case for 'poor' health states, where subjects may not wish to extend life, and for long durations of time where time preference becomes a significant factor. The assumption of linearity in utility for duration intuitively does not hold for states of severe health where increasing duration in a poor health state may be valued less highly than a shorter duration of life in a (marginally) better health state. Sutherland et al (1982) explored this relationship using a certainty equivalence method and found that as duration of survival increased or the health state being valued became more severe, more subjects preferred immediate death to survival in a state of very poor health. This is the concept of 'maximal endurable time', which is dependent on the state of health under consideration.

Bleichrodt (2002) describes other biases that affect the valuation processes of SG and TTO in addition to the effect that time preference has on the shape of individual indifference curves. Probability weighting affects valuations obtained by SG, in that subjects evaluate probability as an ‘s’-shaped function (and not as a linear function), under-weighting small probabilities and over-weighting large probabilities. SG valuations will reflect this weighting, whereas TTO valuations, which are elicited under conditions of certainty, do not.

Loss aversion - the tendency to value losses more highly than equally-sized gains (Kahneman and Tversky 1991) - has been demonstrated to exert important effects on both SG and TTO valuations, suggesting that the starting or 'reference' point for the valuation is particularly important. The reluctance to sacrifice length of life in order to improve health in the conventional TTO (CVG) question (an upward bias in valuation) should be reversed if the subject is asked a ‘reverse’ of the TTO question (CVL) i.e., what number of years in a poor health state is equivalent to a fixed number of years in full health.

Spencer (2003) and Clarke et al (2003) have investigated the effects of altering the reference point of the subject on the valuations of health states. Spencer (2003) investigated valuation using an ‘unconventional’ TTO approach where subjects were asked to imagine living in a health state (less than full health) for a period of two years, followed by death. They were then asked to trade by prolonging life in a lower quality health state. The health states evaluated in this way were EQ-5D states 21222 and 22232. Figure 4 shows that this is effectively, in terms of the theoretical framework introduced in Section 1.1, a compensating variation for a health loss (CVL).

Figure 4. Spencer’s (2003) ‘unconventional TTO’.



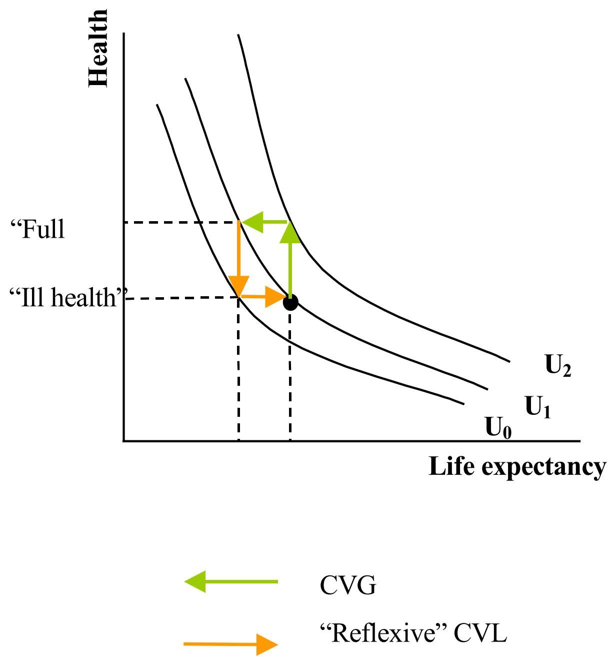
The values obtained for the milder of these two states using Spencer's 'unconventional' approach were significantly lower than the values obtained in the conventional (CVG) manner. However the valuations produced may have been affected by the short duration of time used in this TTO experiment. Short time periods have been shown to affect the willingness of subjects to trade time for quality of life, although most studies reported using short durations relate to palliative care for cancer, e.g., McNeil et al (1981).

Clarke et al (2003) also consider the effect of the valuation reference point and describe the 'unconventional' TTO as a time gain, where subjects start in full health with an endowment of time and gain time in order to compensate for a reduction in health. Again, this is, in our theoretical framework, a CVL. Valuations were elicited by questionnaire using EQ-5D health states 11122 and 22222 and the amounts of time that could be traded were pre-specified. Results showed that subjects may produce inconsistent valuations when asked to value states from different reference points.

Bleichrodt et al (2003) also found violations of consistency at low durations when testing an alternative version of TTO. This study consisted of two stages. Stage one was a conventional TTO (CVG in our taxonomy) in which life expectancy was exchanged for an improvement in health. Stage 2 was 'reflexive', in that respondents were asked to consider starting in full health and with the reduced life expectancy determined in stage 1 and were asked what extension in life expectancy they would require as compensation for a reduction in health to the level of ill health specified in stage 1 (see Figure 5). The elicitation of values in each stage was separated by a gap of at least two weeks. In theory, the subjects should return to the duration in the state of poor health presented in the first question since the individual biases referred to above should apply in reverse when the question is reversed, i.e. the valuations should be 'reflexive'. In particular, when compared with SG valuations, the inconsistencies found at lower durations could be predicted by loss aversion. Longer durations within the questions elicited TTO valuations from both types of question that exceeded SG valuations for the same health states.

Thus, while a number of recent contributions to the literature have considered alternative formulations of the TTO procedure, most - including Spencer (2003), Clarke (2003) and Bleichrodt (2003) constitute use of a CVL. As far as we can determine, none of the existing contributions to the literature recognise the CVG and the CVL as part of a wider class of valuation approaches suggested by Hicks utility theory, and none propose or investigate the equally plausible set of approaches suggested by the EV

Figure 5. Bleichrodt's (2003) "reflexive" TTO.



Existing research does, however, suggest a number of factors to be considered in empirical testing of the four approaches described in Section 1.1. First, the importance of time preference in TTO valuations is expected to exert an influence on the valuation for a given state generated using either CVG and EVL – where the values the numeraire (life expectancy) can take are effectively "bounded" by the initial "reference point" life expectancy against which tradeoffs are made – compared to the valuations for that same state generated by either the CVL or EVG, where the numeraire is unbounded. Second, and also in relation to the "unboundedness" of the numeraire in CVL and EVG, Sutherland's (1982) concept of maximal endurable time may have implications for the nature of the preferences that are evident by testing and comparing all four approaches. Third, prospect theory suggests that gains and losses may be valued asymmetrically, suggesting the value assigned to a given state may be influenced by whether that state is framed as a gain or a loss from an initial reference point.

2 The pilot study

2.1. Aims of the pilot study

From the theoretical framework described above, a pilot study was developed with the following aims:

(a) To investigate the feasibility of valuing health states using EV approaches.

(b) To investigate whether the valuations for a set of health states differ statistically between the four valuation approaches.

(c) To explore intra-rater variation in health state valuations between each approach

(d) To generate insights, via both quantitative and qualitative analysis to facilitate further research.

2.2. Pilot methods and sample

An interview protocol was developed to elicit valuations for three EQ-5D[i] health states: 22222 ("some problems" with mobility, self-care, usual activities, pain/discomfort and anxiety/depression respectively), 11211 ("some problems" with usual activities; no problems on any of the other dimensions) and 11121 ("some problems" with pain/discomfort; no problems on any of the other dimensions). EQ-5D states were chosen to facilitate comparison of our results with those from the MVH study (Dolan 1996), TTO values from which are widely used in economic evaluation in the UK. Mild states were selected in the light of Sutherland's (1982) findings and also because, given the necessity for us to develop fundamentally new valuation 'props' for this study, we wished to avoid the possibility of states being evaluated as being 'worse than dead' which, in the TTO, requires the interviewer to administer a separate valuation procedure (Drummond et al 1997). Questions were developed to address both CV and EV variants of TTO for both health gain and loss. Pre-testing of the EV questions was undertaken on a small sample of students at the University of Otago. This led us to include the following question in the pilot study, in order to investigate the issue of constant proportionality and monotonicity in the relationship between time and utility:

> *'Can you imagine a state of poor health such that, although you would choose to live in that state of health for a short period of time (say one year) rather than die immediately, you would prefer not to live in such a state of health for a long time (say 20 years)?'*

The valuation of health states in response to TTO is made by calculating x/t at the point of indifference, where x is the length of time in full health and t is the length of time in the state for which the valuation is sought. In the pilot interview the four approaches to TTO question ask the subject to vary *x* and *t* as shown in Table 1.

Table 1. Variation of *x* or *t* by TTO approach and calculation for *x/t* where the fixed value is 10 years.

	x	t	x/t
CVG	variable	fixed	x/10
CVL	fixed	variable	10/t
EVG	fixed	variable	10/t
EVL	variable	fixed	x/10

2.3 Interview design and visual aids

A single TTO board, based on that used in the MVH study but with additional features, was developed such that all four valuation approaches could be demonstrated. Particular attention was given to the wording of questions. EV considers equivalent moves to a new level of utility. Linguistically and logically this is represented by an ‘OR’. Thus, for example, in framing an EVG question, we ask questions such as, ‘You have the choice of an improvement in health OR an improvement in length of life, which would you choose?’; length of life being varied to determine equivalence. In contrast, CV considers changes that involve a return to the original level of utility. Linguistically and logically this is represented by an ‘AND’ statement. Thus, for example, in framing a CVG question we ask questions such as, ‘You face the possibility of an improvement in health AND a reduction in length of life; how much of a reduction in length of life would you be prepared to accept?'

Each interview started with a brief introduction, describing the type of questions that would be included. Anonymity was assured and consent to tape the conversation, in order to obtain qualitative data, was obtained. The valuation exercise commenced with the question described above. Subjects were then asked to rank the three EQ-5D states to be evaluated, in order to check their understanding of the descriptions and to provide a means of checking the consistency of their subsequent valuations with these rankings. Each of the three states was then valued using each of the four TTO approaches, yielding 12 valuation tasks per interview. To avoid ordering effects the tasks were ordered by random draw for each interview.

A convenience sample of 50 subjects was recruited from students at City University, staff at a pharmaceutical company and personal contacts. Subjects were also asked to provide some demographic data, (Table 2). Due to its nature the sample is unrepresentative of the general population in terms of age and education. This should be borne in mind in the interpretation of the results.

3 Pilot study results

Demographic characteristics of the participants are shown in Table 2.

Table 2. Characteristics of the respondents

	Number	Percentage
Age		
< 20	0	0%
20-25	3	6%
26-30	8	16%
31-35	17	34%
36-40	9	18%
41-50	5	10%
51-60	1	2%
61-70	2	4%
71-80	0	0%
> 80	5	10%
Sex		
Male	25	50%
Female	25	50%
Educational attainment		
Minimum secondary education	3	6%
All secondary education	7	14%
First degree	16	32%
Postgraduate degree/ Professional Qualification	24	48%
Experience of serious illness in last two years		
Yes	22	44%
No	28	56%

3.1 Feasibility

The order of questions in each interview was chosen by random draw. In practice some question formats were harder for subjects to understand than others, particularly where the initial question was in the CVL format when compared to an ordering where this type of question was asked for the first time later in the interview. The easiest formats appeared to be the CVG and the EVL. Both approaches have a similar output: a trade of a reduction in time to achieve full health, but whereas the CVG finds the change in duration required to return subjects to their initial utility level, EVL seeks the change in life expectancy required to achieve a new utility level. Subjects' comments on the tasks, and the differences in their valuation of the same states between valuation approaches, suggest that subjects generally were conscious of differences between the tasks they were asked to complete.

3.1.1. Non-trading responses

Two types of non-trading response to TTO questions were observed during the interviews:
(a) A refusal to decrease time in full health in the CVG and EVL questions; and
(b) A refusal to increase the number of years in a poor health state in the CVL and EVG questions.

The conventional interpretation of the former case, is that subjects do not trade are effectively saying that the health state being valued is equivalent to full health. In the analysis this type of response is included as a value = 1. In the latter case subjects are stating that *no* increase in time could compensate for, or be considered equivalent to, the change respectively to (CVL) or from (EVG) the poor state of health; for the purposes of the following analyses, these were assigned a value = 0[ii]. This latter type of non-trade was more common in evaluations of state 22222. In the EVG question, where the choice is between increased health *or* increased time, only 20 subjects were prepared to trade time for 22222, with 30 refusing to trade. Rates of non-trading responses are summarised in Table 3 below.

Table 3. Rates of non-trading response to TTO questions by question type and health state.

Health State		11211	11121	22222
Non-trade type (a): State equivalent to full health				
CVG	number of subjects	21	19	1
	percentage of subjects	42%	38%	2%
EVL	number of subjects	15	16	5
	percentage of subjects	30%	32%	10%
Non-trade type (b): due to perceived severity of state				
CVL	number of subjects	6	4	21
	percentage of subjects	12%	8%	42%
EVG	number of subjects	8	9	30
	percentage of subjects	16%	18%	60%

The rates of both types of non-trading response were affected by the state under valuation, the participant's interpretation of the state and the perception of all three states in the interview.

3.2. Valuation results

3.2.1. Initial question on monotonicity of the utility time relationship

48 of 50 subjects interviewed (96%) answered yes to the question: 'Can you imagine a state of poor health such that, although you would choose to live in that state of health for a short

period of time (say one year) rather than die immediately, you would prefer not to live in such a state of health for a long time (say 20 years)?'

3.2.2. Individual ordering of health states for valuation

In their initial ranking of the three states, all subjects indicated that state 22222 was the least preferred.

Table 4. Initial ranking of health states

Order of states	Number	Percentage
11211, 11121, 22222	28	56%
11121, 11211, 22222	22	44%

This difference in ordering is reflected in the summary statistics, where the mean valuations for these two states are similar.

3.2.3. Consistency in individual ordering of health states

Intra-rater reliability was tested by examining whether the order of the valuations for each state given by individual subjects was consistent with their initial ranking of the states. Subjects initially rating state 11211 > 11121 gave the following order of valuations within the interview (see Table 5).

Table 5. Consistency of valuations with ordering for 28 subjects initially rating state 11211>11121

Method of valuation	Order consistent with initial rating		Valuations equal		Order reversed from initial valuation	
	n	%	n	%	n	%
CVG	10	36%	9	32%	9	32%
CVL	13	46%	6	21%	9	32%
EVG	11	39%	6	21%	11	39%
EVL	17	61%	4	14%	7	25%

Similarly, subjects who initially ranked 11121 > 11211 had levels of consistency as shown in Table 6.

Table 6. Consistency of valuations with ordering for 22 subjects initially rating state 11121>11211

Method of valuation	Order consistent with initial rating		Valuations equal		Order reversed from initial valuation	
	n	%	n	%	n	%
CVG	6	27%	8	36%	8	36%
CVL	11	50%	5	23%	6	27%
EVG	7	32%	8	36%	7	32%
EVL	8	36%	8	36%	6	27%

Many subjects reversed their initial rating of these two states within the main body of the interview; others rated the states as being equal. This, together with the fact that the states were not ordered consistently between the 4 valuation approaches, probably reflects the near equivalence of these two states, rather than indicating difficulties with the interview process itself.

3.3. Summary statistics for the valuations

The distribution of valuations for each health state and each valuation type are shown in the appendix. The effect on the distribution caused by non-traders for the CVL and EVG methods is visible, especially for the state 22222. At least one of the distributions for the valuations of the two mild health states, 11211 and 11121, are highly skewed. Medians and ranges for valuations of each state using each method are provided in Table 7.

Table 7. Median and range of sample valuations for each health state and TTO variant

Health State		11211	11121	22222
CVG	median	0.90	0.90	0.60
	minimum	0.50	0.50	0.20
	maximum	1.00	1.00	1.00
	range	0.50	0.50	0.80
CVL	median	0.77	0.77	0.40
	minimum	0	0	0
	maximum	1.00	1.00	1.00
	range	1.00	1.00	1.00
EVG	median	0.74	0.77	0.00
	minimum	0.00	0.00	0.00
	maximum	1.00	1.00	1.00
	range	1.00	1.00	1.00
EVL	median	0.90	0.90	0.60
	minimum	0.50	0.40	0.20
	maximum	1.00	1.00	1.00
	range	0.50	0.60	0.80

The median valuations are lower for the two mild states when the subjects are asked to vary *t* (CVL and EVG) when compared to questions where *x* is varied (CVG and EVL). In the valuation of the state 22222, the median valuations are higher for the questions where *t* is varied than for those where *x* is varied.

The median valuations for the states 11211 and 11121 are all identical except for those generated by the EVG approach. This would be expected given the results shown above for the group where 28 subjects ranked the state 11211 as more preferred to 11121 and 22 subjects ranked these states in the reverse order.

The mean and standard deviation from the mean for each health state and each type of TTO question are shown in Table 8. Non-trading responses are included in the analysis as described above. All groups therefore contain 50 responses.

Table 8. Mean (standard deviation) and number of observations for each health state by TTO approach.

Health State		11211	11121	22222
CVG	mean *(sd)*	0.88 (*0.15)*	0.86 (*0.16)*	0.60 (*0.19)*
CVL	mean *(sd)*	0.68 (*0.28)*	0.71 (*0.24)*	0.38 (*0.36)*
EVG	mean *(sd)*	0.63 (*0.32*)	0.63 (*0.34*)	0.28 (*0.36*)
EVL	mean *(sd)*	0.86 (*0.14)*	0.82 (*0.18)*	0.61 (*0.21)*

Again the mean valuations for the states 11211 and 11121 are similar, as expected from the ordering of states described above.

The skewed nature of the distribution of the valuations obtained in this study means that non-parametric testing between the groups of valuations is required. The valuations for each variant of the TTO question for a given health state were compared, using a two sample Wilcoxon rank-sum test. Analysis was carried out using Stata version 8.0 (Stata corporation, Texas). The following results were obtained from comparison of the groups.

Table 9. Pair-wise comparison of valuation methods for a given health state.

Health State	Pair-wise comparison	Z statistic	probability
11211	CVG vs. CVL	4.655	<0.001
	CVG vs. EVG	4.825	<0.001
	CVG vs. EVL	0.977	0.33
	CVL vs. EVG	0.701	0.48
	CVL vs. EVL	-3.888	<0.001
	EVG vs. EVL	-4.230	<0.001
11121	CVG vs. CVL	3.868	<0.001
	CVG vs. EVG	4.041	<0.001
	CVG vs. EVL	1.031	0.30
	CVL vs. EVG	0.430	0.66
	CVL vs. EVL	-2.748	0.006
	EVG vs. EVL	-3.057	0.002
22222	CVG vs. CVL	2.897	0.004
	CVG vs. EVG	4.248	<0.001
	CVG vs. EVL	0.059	0.95
	CVL vs. EVG	1.385	0.17
	CVL vs. EVL	-2.949	0.003
	EVG vs. EVL	1.964	<0.001

All comparisons except those between the CVG and EVL methods and between the CVL and EVG methods were significant at the 5% level. These are the results predicted from the utility map (Figures 2a and 2b), since valuations made by the CVG method cover the same utility space as those conducted by EVL. The same is true for valuations conducted by the CVL and EVG methods.

The conventional (i.e., CVG) TTO approach used in this pilot facilitates comparisons of these valuations with those from the MVH study (Dolan, 1996) – see Table 10.

Table 10: Comparison of study valuations for CVG TTO with those from the MVH tariff.

Health State	CVG Valuation	MVH Tariff Value
11211	0.88	0.84
11121	0.86	0.83
22222	0.60	0.49

4 Discussion

This paper has reported on a preliminary examination of the four alternative ways of valuing the same change in health with the same initial endowment of life expectancy. Given that there are no theoretical grounds for asserting the primacy of one method over another, the choice of technique is an empirical matter and should rest on issues of acceptability, bias and reliability. There was no indication from our discussions with participants that any of the methods was unacceptable, although some commented on the 'oddness' of EVG and CVL questions:

> *'It's a peculiar compensation, isn't it, I've got pain, so I want more life…'*

Participants also had some difficulty when the randomisation process in our research design meant that all or most types of valuation question for the same state were clustered together. For example, one participant, asked the EVL and CVG questions sequentially for the same state, commented:

> *" This is the same question, one year, I've already answered this one"*

Subjects were also concerned with the 'correctness' and consistency of their valuations; some attempted to value the same state consistently between methods.

> *" I'm sure we've already done this question, what did I say last time? I think seven, so seven is the right answer"*

In the absence of a gold standard measure of utility, the question of bias is difficult to determine. Older and younger respondents made comments that indicated their views on time preference were, predictably, rather different and this may have affected the valuations for EVG and CVL, where the equivalent or compensating change in life years in less than full health is unbounded. While, as we have noted, there was a considerable proportion of non-traders on the EVG and CVL, comments from those who did trade made comments that indicate their values were affected by their views about the future:

> *" That would give me an extra 20 years, but it would mean that I would stay in pain, but I think I would rather live longer and hope that there would be more treatment out there that would help my pain in future".*

Reliability might be inferred from the variability in the answers that people gave for the same state of health. The variability of responses, as judged by the standard deviations of the estimated mean values ascribed to the health states, did not differ appreciably between methods.

The most significant finding relates to the use of the approaches that ask people to consider extensions in life expectancy, either as an alternative to a health gain (EVG) or as a compensation for a health loss (CVL). Two key pieces of evidence from our pilot suggest a relationship between length of life and utility that is non proportional and even non monotonic for some states of health.

First, a substantial proportion of our sample was insufficiently attracted by the possibility of longer life in poor health to trade. This was particularly the case with state 22222:

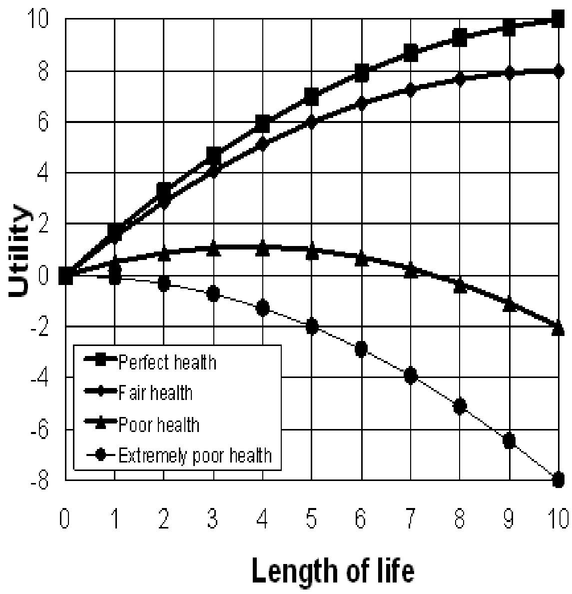
> *"…the thought of going longer in that state is not good"*

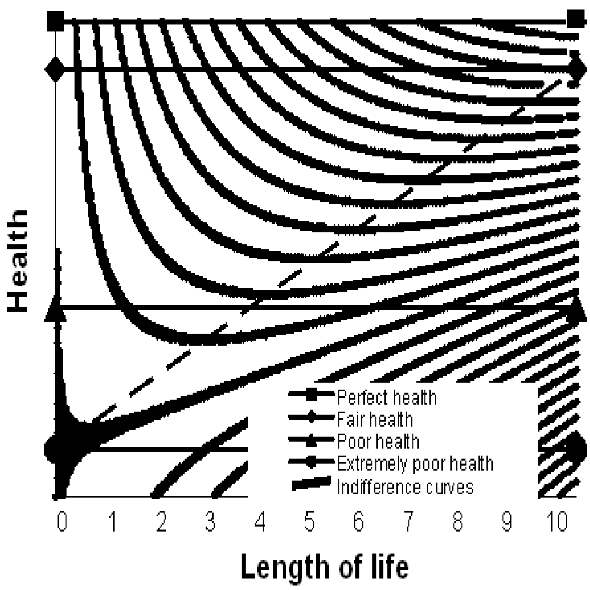
Second, 96% of respondents responded positively when asked "Can you imagine a state of poor health such that, although you would choose to live in that state of health for a short period of time (say one year) rather than die immediately, you would prefer not to live in such a state of health for a long time (say 20 years)?

Taken together, these findings suggest the possibility of utility functions in length of life of the general form illustrated in Figure 6a.

Figure 6a. Utility as a function of length of life

Figure 6b. Indifference curves in health and length of life





In perfect health we assume that utility never declines with increasing length of life. In such states of health, we might like to live forever. At some 'poor' states of health utility increases then decreases. We would like to live for some time in such states, but not forever. Where this is the case, our representation of utility as a function of time is non-monotonic attaining a maximum value (an inverted 'u-shape'). Some 'extremely poor' states of health are shown such that utility invariably decreases with length of life. We would prefer to die immediately rather than endure such states. These states are unambiguously worse than death. Figure 6b illustrates the same relationship in the form of indifference curves.[iii]

Figure 6b helps us to understand and interpret some of the results from the four valuation procedures when considering those poor states of health for which a maximal endurable time exists. Such a utility mapping would have the following implications:

(a) Equivalent variations of a health-related quality of life gain may well be unstable or even indeterminate at low quality of life levels.

(b) There is no single health state that can be described as a ‘state worse than death’. It depends upon the duration of the state. This is illustrated by the line showing, for any given health state, the maximum total utility possible in increasing length of life – the ‘maximal desirable length of life’. Beyond this point, in any given state of health, increased length of life is detrimental to utility.

(c) Although not so unstable as the equivalent variation, compensating variations may also be unstable when considering poor states of health.

These results pose a fundamental challenge to the appropriateness of the constant proportionality assumptions invoked in the estimation of TTO values and their application in CUA. If utility accumulates at different rates as length of life changes, a single value for any given state, independent of its duration, is not tenable.

There is no simple way out of this dilemma. Theoretically, if we knew exactly how utility varied with life expectancy, it might be possible to integrate a time utility function[iv]. At present, when calculating QALYs, we simply ignore the problem. This latter may be a reasonable approximation if we are considering relatively lengthy time periods, and if the anomalies in the time utility are no more than ‘local disturbances’ applying to short time frames, with reasonable stability over longer periods.

If our aim is to produce an index by dividing an estimate of the utility of some combination of health state and length of life, by the time over which that utility is accrued, then it may be preferable to control that length of life in constructing the question. The CVG and the EVL both value a specified life expectancy in poor health and trade time in good health. The CVL and the EVG effectively value a variable time in poor health using a specified time in good health. This implies poor control over the thing being valued. Moreover, the findings from our pilot study suggest that, even in intermediate states of health, it may not be possible to extend the time in poor health sufficiently to achieve a change equal to perfect health. We need to explore whether this problem arises if we start with much shorter life expectancies. It is likely that more people would be content to live longer in poor health if the choice was between that and a very short life in perfect health.

More work is needed to explore these issues. In the meantime, we should not assume that the conventional TTO is applicable to a wide range of life expectancies. It appears that even moderate states of health may be dispreferred to death over long durations.

Acknowledgments

The authors wish to express their thanks to the participants in the pilot study. We are also grateful to Paul Kind, University of York for allowing us access to the TTO board developed and used by the MVH study, and for comments provided by participants at a seminar we gave at Sheffield University.

References

Bleichrodt H. (2002) A new explanation for the difference between time trade off utilities and standard gamble utilities. *Health Economics* 11: 447-56.

Bleichrodt H, Pinto JL. Ebellan-Perinan, J.M. (2003) A consistency test of the time trade-off. *Journal of Health Economics* 22:1037-52.

Buckingham K (1995) *Unifying the theory of 'health gain' valuation*. Paper presented at HESG, Bristol.

Clarke P, Wolstenholme J, Johnston K. (2003) *Time trade-off or time gain? It all depends on your point of reference.* University of Oxford Working Paper. Health Economics Research Centre.

Dolan P. (1997) Modelling valuations for EuroQol health states. *Medical Care* 35: 1095-1108.

Dolan P, Gudex C. (1995) Time preference, duration and health state valuations. *Health Economics* 4:289-925.

Dolan P, Stalmeier P. (2003) The validity of time trade-off values in calculating QALYs: constant proportional time trade-off versus the proportional heuristic. *Journal of Health Economics* 4:445-458.

Drummond MF, O'Brien B, Stoddart GL, Torrance GW. (1997) *Methods for the economic evaluation of health care programmes.* (2nd edition). Oxford Medical Publications.

Hicks JR. (1944) The four consumers' surpluses. *Review of Economic Studies*.

Johannesson M, Jonsson B, Karlsson G. (1996) Outcome measurement in economic evaluation. *Health Economics* 5: 279-96.

Johansson P-O. (1991) *An introduction to modern welfare economics*. Cambridge University Press.

Kahneman D, Tversky A. (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47:263-291.

McNeil BJ, Weichselbaum R, Pauker SG. (1981) Speech and survival: tradeoffs between quantity and quality of life in laryngeal cancer. *New England Journal of Medicine* 305:982-87.

Mishan EJ. (1971) *Cost benefit analysis.* London: Allen and Unwin.

Ng Y-K. (1979) Welfare economics: Introduction and development of basic concepts. London: Macmillan.

Parkin D, Devlin N. (2004) *Is there a case for using Visual Analogue Scale valuations in CUA?* Discussion Paper 04/03, Economics Department, City University, London. http://www.city.ac.uk/economics/discussionpapers.htm

Perez DJ, McGee R. Campbell AV, Christensen EA, Williams S. (1997) A comparison of time trade-off and quality of life measures, *Quality of Life Research.*

Richardson J. (1994) Cost Utility Analysis: what should be measured? *Social Science and Medicine* 39: *pages numbers to be confirmed.*

Shiell, A., King, M. Briggs, A. (1993) The consistency of rating scale and time trade off techniques for eliciting preference weights for health states. Paper presented to the HESG Conference, July.

Spencer A. (2003) The TTO method and procedural invariance. *Health Economics* 12:655-68.

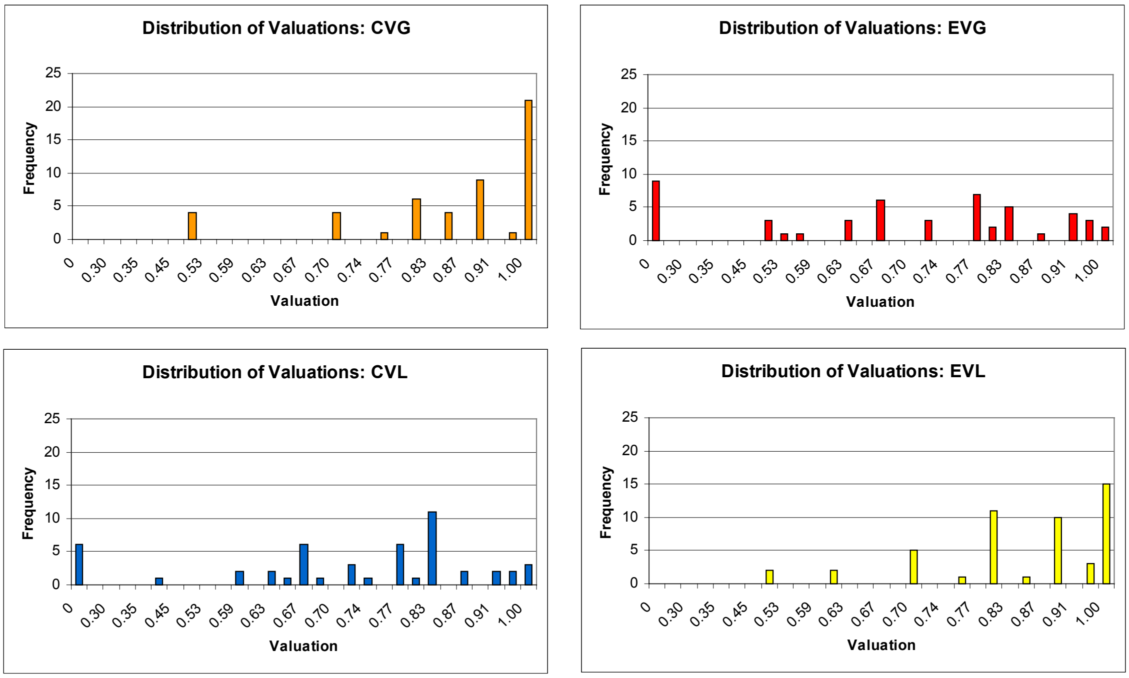
Sutherland HJ, Llewellyn-Thomas H, Boyd NF, Till JE. (1982) Attitudes towards survival. The concept of "maximum endurable time". *Medical Decision Making* 2:299-309.

Torrance GW. (1976) Towards a utility theory foundation for health status index models *Health Sciences Research* Winter: 349-369.

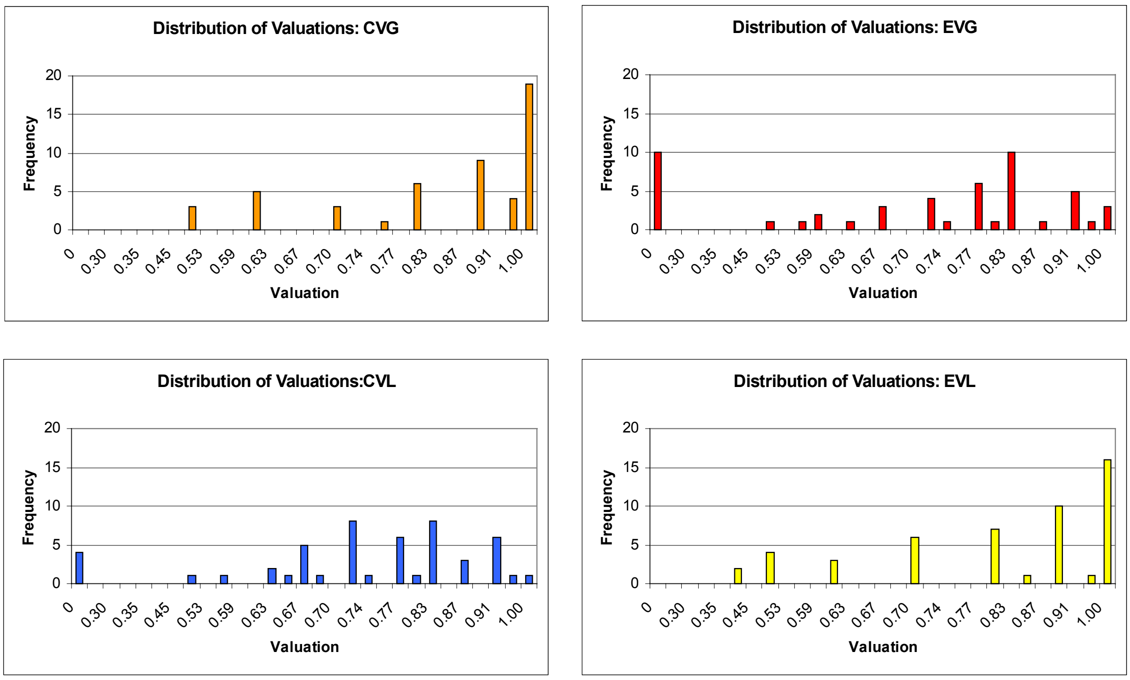
von Neumann J, Morgenstern O. (1944) *Theory of games and economic behaviour*. Princeton University Press.

Appendix

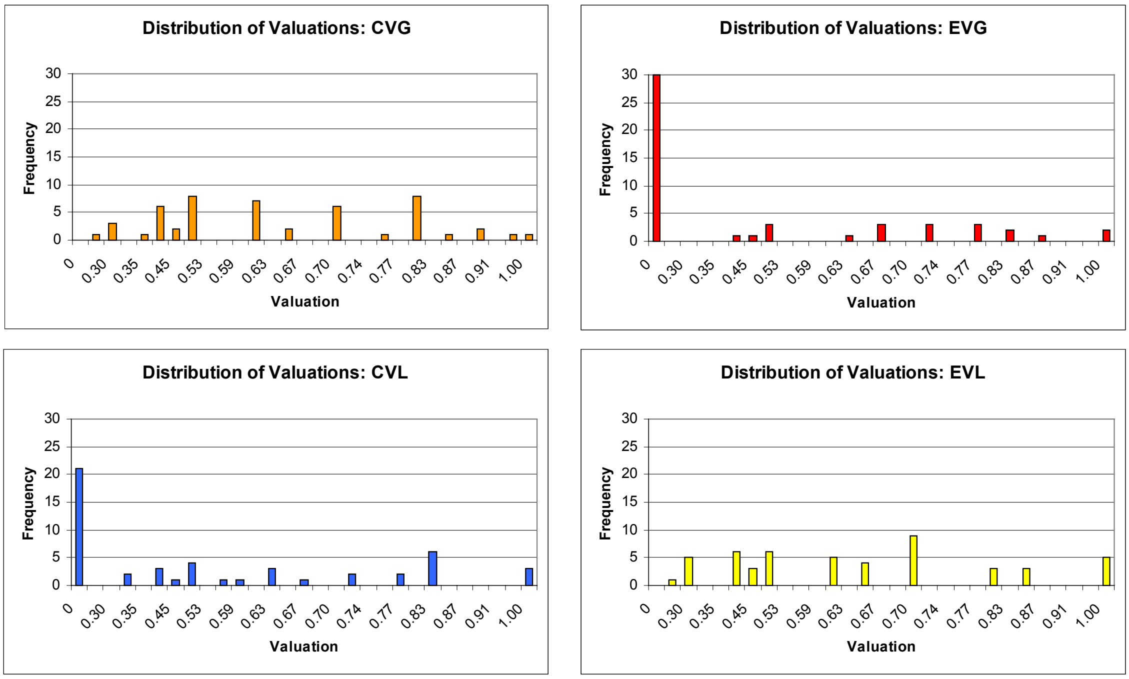
Frequency Distribution for Valuations of Health State 11211



Frequency Distribution for Valuations of Health State 11121



Frequency Distribution for Valuations of Health State 22222



---

[i] The EQ-5D is an instrument for describing health states. Health state are described in terms of 3 levels (‘no’, ‘some’ or ‘extreme’ problems, coded as 1,2 and 3 respectively) on five dimensions (mobility, self-care, usual activities, pain/discomfort and anxiety/depression. The EQ-5D describes 243 unique states, each identified by a five digit number (e.g. 21111) according to the levels within each of the dimensions, in the specific order listed above.

[ii] Including these non-trades at 0 assumes that these states were considered ‘at least as bad as’ dead, given that the scale of values is anchored at dead = 0 and full health = 1. This may impose be an upward bias on mean values for the states concerned, but was considered preferable to the bias introduced by excluding these valuations.

[iii] Figures 6a and 6b were produced using the same parameters in a function of the general form $U = aHL - bL - cL^2$ where U = Utility, H = Health and L = Length of life.

[iv] Alternatively we could use a composite value such as a HYE to aggregate the utility of individual length of life/quality of life combinations. The use of HYEs in this way would, however, require a large research effort to construct a tariff of values for the multiplicity of possible health *profiles* rather than a tariff for health *states*.