Time Trade Off valuation of EQ-5D states worse than dead: a feasibility study

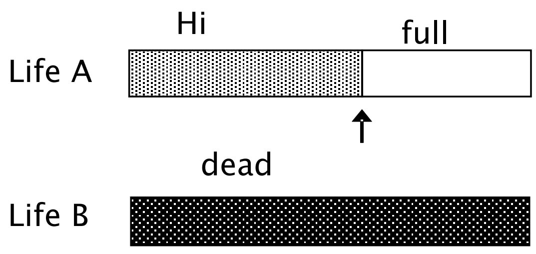
Nancy Devlin[a], Aki Tsuchiya[b,c], Ken Buckingham[d] , Carl Tilling[c].

a. Economics Department, City University, London
b. Department of Economics, University of Sheffield, Sheffield
c. School of Health and Related Research, University of Sheffield, Sheffield.
d. Department of Preventive Medicine, University of Otago

1. Introduction

The protocol for eliciting Time Trade Off (TTO) valuations of EQ-5D health states used in the York Measurement and Valuation of Health (MVH) study (Dolan 1997), although not formally endorsed by the EuroQol Group, has become the Group's *de facto* protocol and is widely used in part or full by other researchers. For example, the MVH protocol was used in research to produce TTO value sets for The Netherlands (Lamers et al 2006) and the US (Shaw et al 2005). The MVH protocol for valuing states worse than dead (swd) requires participants to choose between immediate death, and spending a length of time (10-*x*) in H_i (the state to be valued), followed by *x* years in full health. The value of *x* is varied until the participant is indifferent between the two options. This procedure is illustrated in Box 1. The worse a given health state H_i, the more time is required in full health to compensate for the time spent in H_i. The value for H_i is given by $U(H_i) = -x/(10-x)$. For example, if the participant is indifferent between the option of immediate death, and the option of 6 years in H_i, followed by 4 years in full health, then the value of $U(H_i) = -4/(10-4) = -4/6 = -0.67$.

Box 1. The conventional TTO approach to valuation of states worse than dead.



There are a number of problems with this approach to valuing swd. The trade-off procedure employed for swd is fundamentally different, both conceptually and operationally, from that used for states *better* than dead (sbd). Conceptually, the TTO for sbd involves finding the number of years (*x*) in full health that is equivalent to *t* years in H_i, at which point $U(H_i) = (x/t)$. In this case, the denominator *t* is 'fixed' (in the MVH protocol, to 10 years) and the value is obtained by varying *x*. In contrast, the TTO procedure for swd involves *simultaneously* changing both the numerator and the denominator. Operationally, the use of different elicitation procedures for sbd and swd means that at the point during the elicitation procedure where it becomes evident the valuation for a given state is negative, the interviewer has to switch to the use of a completely different 'prop' and the participant has to re-engage with an entirely different sort of task. Given these differences, and other problems discussed below, the aggregation of positive and negative values for any given state in order to calculate mean values or to estimate a value set is of questionable validity.

Further, the MVH elicitation procedure produces ‘extreme’ negative values; and how *much* negative valuations may become depends on the (relatively arbitrary, and seemingly innocuous) researcher decision about the ‘units’ of time that can be traded. In the MVH study participants could make tradeoffs in units of 3 months (0.25 years); hence the worst possible health state had a value of -9.75/(10-9.75) = -39. If the trade-off was in whole years then the minimum value would be -9/(10-9) = -9; or, if it was in months (0.08 years) then the minimum value would be -9.92/(10-9.92) = -124, and so on.

Because the elicitation procedure produces extreme negative values, researchers have responded by transforming or bounding negative valuations to -1 in various ways (see Lamers 2007). Patrick et al (1994) point out that, once transformed, the negative numbers for swd can no longer be interpreted as “utility” scores, measured on the same scale as those for sbd. Yet standard practice in calculating QALYs is to treat all values reported in value sets as commensurable. For example, an improvement from -0.2 (a swd) to 0, experienced over one year, produces a gain of 0.2 QALYs, and this is treated in technology appraisals as identical to an improvement from 0 to 0.2 experienced for one year - whereas the underlying ‘untransformed value’ for the swd might suggest these two improvements in health are valued quite differently. This has serious implications for organizations such as NICE that routinely use the MVH TTO value sets in technology appraisals and decision making.

It has also been observed that there is a ‘gap effect’ in the MVH TTO values around dead. In an analysis of differences in the values for adjacent states, Stalmeier et al (2005) find the differences in TTO values for states either just above or below 0 are at least twice as large compared to the differences between other adjacent states. Further, there are few states in the MVH TTO value that are close to 0 (either sbd or swd). This discontinuity means that, for example, states worse than dead tend to be *substantially* worse than dead. This ‘gap’ is probably related to the separate valuation procedure used for swd. For example, it could caused by the explicit questions asked of participants regarding whether the state in question is indeed better or worse than dead. The notion of a swd may be disturbing to members of the public, and having declared a state as being so, there may be a focusing effect which amplifies how *much* worse the state is than dead.

The problem of TTO valuations of swd is not unique to the EQ-5D. Other research groups that use the TTO face the same issues, although the response in each case – both the specific elicitation procedures employed, and *post-hoc* manipulations of the data – differ. Neither is the problem restricted to TTO. The same problem of extreme negative values also affects SG and VAS, and should researchers choose to transform the negative values to be bounded by -1 then, similarly, the resulting values will be incommensurable with values for sbd.

A crucial issue is whether or not values of negative states *should* be bounded to -1. It is not obvious why there should be *no* states worse than -1. For example, the phrase ‘it would have been better if he had never been born’ could truly be applied to people who have undergone torture and other types of extreme suffering. There is no theoretical basis for imposing a limit on the level of disutility associated with these extreme sufferings.

Box 2. Combining sbd and swd in the estimation of QALYs.

The following combinations of utility (sbd) and disutility (swd) are <u>all</u> equal to 50 quality adjusted life years (QALYs), (assuming a zero rate of time preference).

Live to 50 in full health and die; OR
Live to 51 in full health, then live in -1 for 1 year, and die at 52; OR
Live to 52 in full health, then live in -2 for 1 year, and die at 53; OR
Live to 53 in full health, then live in -3 for 1 year, and die at 54.

If, as is the case, we interpret sbd as states that contribute positively to one's stock of lifetime QALYs, then we may correspondingly interpret swd as those that contribute *negatively* to one's stock of lifetime QALYs. If so, then all the scenarios in Box 2 are equivalent in QALY terms. This suggests that values less than -1 may be both conceptually plausible and represent meaningful expressions of value. If this point is conceded, then the justification for transformation is greatly reduced - while increasing the importance of designing a means of eliciting negative valuations that are meaningful.

This study aims to develop and test a TTO protocol that overcomes the issues with the MVH approach, and to explore the characteristics of the valuation data it generates. Specifically, our objectives are to (a) identify and compare the various TTO valuation procedures employed by researchers in the valuation of swd; (b) propose a valuation procedure which overcomes the problems encountered with the MVH protocol; (c) pilot, refine and test the feasibility of using the proposed protocol, and (d) report the resulting valuation data. In what follows, section 2 summarises a literature review on TTO designs, sections 3 and 4 are on the methods and results of the pilot study, sections 5 and 6 present the methods and results of the feasibility study, and section 7 concludes.

2. Methods for valuing states worse than dead

Our search for alternative TTO procedures included both a review of the empirical literature and a conceptual analysis.

We undertook a literature review of all studies which had elicited TTO values, to identify how swd had been handled. Both are reported in full in Tilling *et al* (forthcoming). TTO was predominantly used to elicit valuations from patients. In almost all these cases, no protocol for swd is mentioned, reflecting practical and ethical constraints.

Of the 375 papers identified, less than 10% of the papers reporting TTO elicitation mention swd; of these, four distinct approaches were identified - the MVH approach; Torrance (1982), Detsky et al (1986) and Robinson and Spencer (2006). Stalmeier et al (2007), although not employing the TTO, was also considered relevant. Each approach was assessed on the following criteria: uniformity of procedure for swd and sbd; minimising cognitive burden on participants; participant comprehension of the task; concordance of values with patients' views about health; minimising bias.

Torrance's TTO procedure for swd is very similar to the MVH, except that the order of full health and H_i are reversed (i.e. in Life A, from full health to H_i, rather than H_i to full health as in Box 1). The Torrance approach was rejected on the grounds that it suffers the same problems as the MVH protocol.

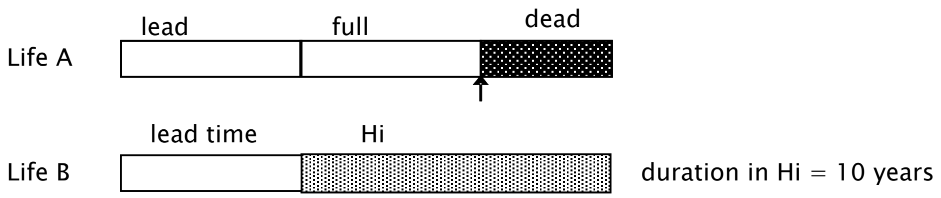
Detsky et al. (1986) report a two-stage TTO procedure with chaining. Participants first value a sbd in the usual fashion. They are then asked to choose between one month in that state and one month comprising a number of days in a second health state (which may be a swd) and the remainder in full health. The number of days in full health is varied until the participant is indifferent between the two options. The point at which indifference is achieved is chained via the previously obtained value for the sbd to produce a value for the swd. This approach has several disadvantages. It is time consuming and complex and cannot calculate the value for a swd unless a sbd has already been valued. In addition, the second stage of the TTO implies variable denominators, and consequently values for swd may become extremely negative, as in the Torrance and MVH designs. The paper does not report whether there were values less than -1, and if there were, what was done about them.

Stalmeier et al (2007) asks participants to make a binary choice between whether they think a health state is better or worse than dead, and these responses are used to estimate values at the aggregate levels. If state H_i is preferred to being dead (by an individual) it receives a score of 1; if considered equal to dead, 0; and if considered worse than dead, -1. Scores are then aggregated across respondents. For example, if 70%, 20% and 10% of participants had respectively responded with 'better than dead,' 'equal to dead' and 'worse than dead', then the value would be: (0.7×1) + (0.2×0) + (0.1×-1) = 0.6. While an interesting approach, there are a number of weaknesses. For example, suppose there was a health state that everybody would value at 0.5. Using this approach the state would receive a value of 1 rather than 0.5.

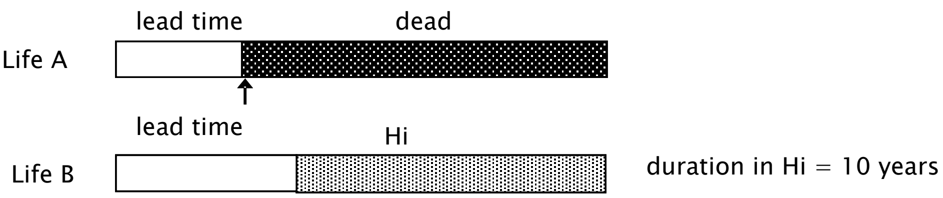
Spencer and Robinson (2006) report a single TTO procedure that can be used for swd and sbd. Participants ranked a series of life profiles consisting of combinations of good health, poor health and death. The principal innovation of the method is its introduction of a 'lead time' in full health preceding each of the alternatives presented – see Box 3. The approach avoids the need to have different valuation procedures for sbd and swd by allowing participants to trade their lead time to avoid states considered worse than dead. We considered this the most promising of the alternatives we had discovered, and selected it as the basis for our study.

Box 3. The Lead Time TTO, with an illustration of responses that would yield positive and negative valuations respectively.

State happens to be 'better than dead':



State happens to be 'worse than dead':



We also undertook a conceptual analysis to establish the wider set of TTO methodologies that might be constructed using the key components of currently applied TTO methods – in order to determine where there might be a variant which had not been subject to empirical analysis. The components are: time spent in the health state to be valued; time spent in perfect health; and time spent dead; these being combined into two options with an equilibrium achieved by varying time spent under one of the options, the latter referred to as the equilibriating option. All possible combinations of lead times (times in a given state of health common to both options at the start of each), lag times (times common to both options at the conclusion of each), and the states forming the equilibriating option were examined We confirmed that no untried TTO methodologies were feasible using adaptations of existing TTO components. Our analysis did not extend to a consideration of 'chained' designs. For a full report, seeTilling et al (forthcoming).

3. Pilot study design

An interview protocol and physical prop was developed to implement the lead time concept, but mimicking the MVH design i.e. using a TTO board with a sliding device, and similar colours to represent the states. Because we were concerned that the introduction of a lead time potentially implies a greater confounding effect of time preference, we decided to use the pilot study to test and compare two contrasting duration versions of the lead time TTO: one with a short duration (10 weeks) for state H_i, and one with a longer duration (10 years) for comparability with the MVH. The 'lead time' was 10 weeks and 10 years respectively.

The pilot employed three elicitation procedures: lead time TTO (long); lead time TTO (short) and MVH, for three EQ-5D states: 22222, 33333 and 23323. The order of the TTO procedures was varied between interviews so that approximately half of the interviewees encountered the short duration lead time TTO first, and half the long duration lead time TTO. The long duration lead-time TTO was paired with the MVH in this process, so that the two 10-year duration TTO procedures were always undertaken sequentially.

We anticipated a possibility that some participants' distaste of the severe states might be such that they 'traded away' all their 10 year (or 10 week) lead time. We therefore designed the prop so that, should this arise, an alternative TTO board fascia could be superimposed, extending the time scale and lead time to 20 years (or weeks).

Prior to the TTO tasks, participants ranked the states, and provided visual analogue scale valuations of them. Following completion of the TTO valuation tasks, participants were taken through a process of guided reflection on their valuations and their feedback on the tasks - including explicit comparisons between the three protocols - sought using a series of open-ended questions. These included questions to gauge participants' views on the ordering of the states in Life A. In the MVH protocol for swd, Life A comprises time in poor health followed by time in full health, whereas the lead time TTO presents time in full health followed by time in poor health. We wished to explore whether participants had any issues with the *plausibility* of going from extremely poor health to full health in the MVH, and whether their evaluation of 'good years followed by bad' in the lead time TTO had led them to consider the possibility of enjoying the good years, then committing suicide to avoid the later years of extreme misery[1].

Interviews were digitally recorded with participants' consent, and their verbal responses to the elicitation procedures also noted by the interviewers. Interviews were undertaken by KB and ND in London, and by CT in Sheffield. Further details of the physical prop, the interview scripts and booklets are available on request to the authors.

In order to calculate the values from the lead time TTO responses, the lead time is simply subtracted from both the numerator and the denominator to give a result comparable with the usual TTO. Where the lead time is 10 years (or weeks), $U(H_i) = (x - 10)/20\text{-}10$). If at the point of indifference, 'x' years in full health in Life A is *greater* than the lead time in Life B (see Box 3), the value will be positive; if it is equal the value is equal to 0, and if x is *less* than the lead time in Life B, the value will be negative.

4. Pilot study results

A total of 27 participants, ranging in age from 18 to 63, undertook the TTO tasks and provided feedback on them. In subsequent questioning, the aspect of the tasks most frequently reported by participants to be difficult was imagining very poor health states hypothetical to them, rather than concerning the elicitation procedures *per se*.

With respect to the 'bad to good' ordering of the states in the MVH, reactions were mixed. Some participants simply responded that they didn't think it was a problem, or that it hadn't bothered them. A number of respondents stated that they had thought it was unrealistic, but 'went along with it'. With

[1] It was this possibility that led to the MVH protocol reversing the order of full health to poor health in the original approach to valuing states worse than dead described by Torrance (1982).

respect to the ‘good to bad’ ordering in the lead time TTO, 22 participants said they had not thought of the possibility of committing suicide after the good years; five indicated it had occurred to them. Of these, three said that although they had thought of it, they didn’t take it into account because they assumed that would be cheating, and that the experience of poor health was ‘part of the package’ they were asked to consider.

Both short and long duration variants of the lead-time TTO were equally feasible for the participants to complete, judged in terms of participants’ reported ease of completion and understanding, and in terms of the valuation data obtained. However, the contrasting durations provoked considerable reaction by participants regardless of the order in which they were encountered. We expected that the shorter duration of severe states would result in less time being traded and therefore higher values. Some participants’ responses were compatible with this, and responses to the question “Did the change in time scale make a difference?” included: “I would be prepared to put up with more of being poorly because more time is of the essence when you’re thinking you’ve only got 20 weeks” and “Definitely, it’s harder to endure poor health for a long time so death is more preferable”.

In contrast, other participants commented that thinking in weeks somehow made the states ‘more real’, for example: “When its years, I thought ‘oh, I can cope with that’, but when it comes to weeks, if I only had 10 weeks and if I was going to be depressed for 5 weeks, I’d rather just go out with a bang than put up with that. I think that your body can adjust to pain and that over a long period of time”. And in other cases, ‘stage of life’ considerations seemed to outweigh both time preference and maximal endurable time: “Definitely. In weeks…I wouldn’t meet anyone new or get involved in anything new, its just more weeks like what they are now. But 10 years, anything could happen. Like, if I had a partner and a child, I’d probably want to be spending as much time with them as I could?”

The contrasting durations meant our pilot data facilitated a rudimentary test of constant proportionality. Our 27 participants valued three states using two contrasting duration lead time TTOs. Of the 81 opportunities to satisfy it, constant proportionality was satisfied on no occasion. Of the 27 participants, 11 consistently had higher TTO values in the short duration than the long duration; 10 had no consistent pattern of differences between durations; whereas 6 had consistently *lower* TTO values in the short duration than the long duration.

Notwithstanding the considerable variation between participants’ reaction to the different TTO procedures, at the aggregate level much of this seemed to wash out: the mean and standard deviation of values for each state were remarkably consistent across all three methods. For the MVH, the 10-year lead time and the 10-week lead time variants respectively, the mean values for 22222 were 0.62, 0.52 and 0.48; for 33333 the mean values were -0.28, -0.47 and -0.38, and for 23323 the values were 0.04, -0.09 and 0.09. (Note: these mean MVH values are the unbounded values elicited from our sample).

Seven of the 27 participants traded away all of their 10 year lead time when evaluating state 33333 and therefore proceeded to the extended lead time version. In one case, the 20 year lead time was then also exhausted. Correspondingly, four of the 27 participants exhausted the 10 week lead time when evaluating 33333, and in two of these cases, also then traded away all of the extended 20 week lead time. One person also exhausted their 10 week lead time for state 23323.

5. Feasibility study design

Results from the pilot suggested both short and long duration lead time TTOs were equally feasible. The long duration protocol was selected for the feasibility study to enable comparability with MVH results. Furthermore, we decided to field a single protocol only (i.e. not attempting within- or between-respondent comparisons between the lead time and the MVH approaches) as this enabled us to generate more valuation data from the protocol of choice, than to share the limited number of total interview person-time across multiple protocols. This is compatible with the aims of our study, which were to develop and test the feasibility of a uniform procedure for swd and sbd; and to provide insights into the characteristics of the valuation data generated.

Each of the 100 respondents was asked to value 10 EQ-5D states. For comparison, the number of states per participant included in EQ-5D TTO valuation studies was nine in Zimbabwe; 13 in the UK, US, Spanish and German TTO studies; 16 in the Denmark study; 16 and 17 in the Japanese and Dutch studies (Szende et al 2007). The 10 states were selected from the set of states for which valuations were elicited in the MVH study; but drawing principally on severe states where we were more likely to observe responses indicating swd.

The interview protocol and script used in the pilot was refined and improved eg. by making the ‘branching’ instructions clearer. Using the responses to the open ended feedback questions in the pilot as a guide to the principal topics and themes, we developed a series of structured-response questions to capture participants’ views about the elicitation task. We also attempted to gauge each participant’s rate of time preference by asking a short series of questions that involved weighing up the duration and timing of headache episodes.

Given our experience in the pilot with participants ‘using up’ the lead time, and in order to minimise the complexity of using different time scales as in the pilot, we extended the lead time to 15 years while retaining the 10 year duration of the state to be valued. The values were therefore calculated as $U(Hi) = (x\text{-}15)/(25\text{-}15)$. In instances where the 15 year lead time was exhausted, participants were asked a simple yes/no question about whether they would still prefer life A (in effect, as the lead time has been traded away at that point, immediate death) if Life B comprised 20 years in full health and 5 years in *Hi*. If they preferred Life B, their value was set at -1.5. If they still preferred life A then we coded these as (0-20)/(25-20) = -4 as the baseline presentation. However, since one may question the reliability of this last response, which is not based on a purpose built visual aid, an alternative coding gives all responses where the 15 year lead time was exhausted an artificial value of -1.5. Furthermore, in order to facilitate comparisons with the MVH values (which are transformed to -1), a third set of TTO values were generated by dividing all negative values under the baseline approach by 4, so that the minimum value becomes -1.

Interviews were conducted by a team of four interviewers at Sheffield Hallam University (SHU). The interviewers had previous experience in eliciting TTO values, and were given training and practice in using the lead time TTO prop and interview script.

The general public sample was elected using SHU’s standard recruitment frame, “AFD software names and numbers”. A random sample of areas was chosen, based on postcode, in the South Yorkshire area. A letter inviting participation and information sheet on the project was mailed out to addresses in those areas in February 2008. Interviewers then called on the houses in those areas in the following four weeks.

6. Feasibility study results

6.1 Respondents

Valuation and background data are available from 108 members of general public who participated in interviews. Table 1 shows the background characteristics of our sample, with the corresponding values from the MVH study included for comparative purposes. Compared with the MVH sample, our sample contains proportionally slightly more females, fewer people over the age of 64, more people between the ages of 45 and 65, more employed people and fewer retired people. Furthermore, a smaller proportion of our sample had experience of serious illness both in themselves and in their family, but a higher proportion had experienced serious illness in caring for others. Mean self-reported health through VAS is slightly lower amongst our sample, despite the fact that a larger proportion of our sample are employed/self-employed. The net household income categories were based on income quintiles from the Office of National Statistics for 2006-2007. One can see that the category £20,001 - £26,000 is highly under-represented, while the category £26,000 -£40,000 is highly over-represented. Given our small sample size and the opportunistic nature of our data collection it is unsurprising that we have not gained a representative income distribution. Our sample has a very high level of home ownership. The areas in which the respondents were recruited from were coincidentally all dominated by mortgaged properties rather than rentals.

6.2 Ranking, VAS, and TTO main results

It was identified that one of the interviewers misunderstood the ranking procedure, where EQ-5D states were presented in three batches. Instead of ranking the second and third batches of states alongside the first batch of ranked states, this interviewer conducted three separate ranking exercises. This affected 17 respondents; these respondents were dropped from the ranking data, leaving 92 respondents. The results are summarised in Table 2. Average rank gives the rank for each state assuming rank scores are cardinal. State 11111 has the highest and Dead the lowest average rank: no states are, on average, ranked as worse than dead, including the worst possible state: 33333. 'Predicted % 1st rank' gives the predicted probability of a given state being ranked first, based on the rank ordered logit regression results. There is high consistency across the different approaches to summarise rank data.

Table 3 shows both the raw and re-scaled VAS values for both our study and the MVH study. The raw values represent the position that respondents placed a given state on the scale, while the re-scaled values set state 11111 (full health) equal to 1 and dead equal to 0 and the other states are then adjusted depending on their position relative to these two states. The VAS values in our study are consistently higher than those in the MVH study. Furthermore, none of our mean re-scaled values are negative, while two states receive a negative value in the MVH study (33323 and 33333).

Table 4 shows the TTO valuations for the ten health states from our feasibility study, both for the baseline case (with minimum value of -4); when minimum values are restricted to -1.5; and when the negative values are re-scaled from those in which minima were -4 to values in which minima were -1. These are shown alongside the corresponding MVH values (also transformed according to the MVH protocol to give a minimum of -1 for negative values). We would caution the reader against reading too much into these values. Our study was not designed to elicit a valuation set. In particular, we have not sought precise valuations from those people for whom even zero time in good health was preferable to 15 years lead time in good heath with the knowledge that this had to be accepted together with 10 years in poor health. Rather we have simply made various assumptions about the minimum values that their answers imply. Given this important caveat, our results are what might be expected. For the more severe states of health, lead time TTO values are lowest, when minima of -4 are used in the baseline case, and become progressively higher when the minima are -1.5 and -1. In the latter case they generally exceed the values derived for the MVH study. Except for the mildest state, 11112, the mean MVH values are lower than the lower 95% confidence interval for our results with minima -1.

Figure 1 shows the distribution of TTO values in our study compared to those from the MVH Study. There is a large gap from -1.5 to -4 in our study which is a result of the binary question we asked respondents that exhausted their lead time. What is most striking is the peak that occurs at 0 in the MVH study is not present in our lead time TTO results.

6.3 The effect of background on ranking, VAS and TTO

A cross-correlation matrix of key variables (available from authors) showed that only the correlation coefficient for VAS and ranking exceeded 0.7. Other weak correlations (coefficients of at least 0.2) were between: TTO and VAS scores; TTO and Rank scores; own ill health and having a degree; own ill health and own VAS score; experience of health in others and experience as a carer; being a home owner and income greater than £26,000; being older than 45 and being a mother.

Table 5 shows the coefficients and corresponding t-statistics from an ordered logit regression to show the effect of background characteristics on the rankings. These show that none of the characteristics appear to affect the ranking of a given state (see control variable).

Table 6 shows the coefficients and t-statistics from an OLS regression to show the effects of background characteristics on VAS valuations. The results show that those who have themselves experienced illness or have experience as a carer will give a lower value to a given health state. The results also show that males and homeowners gave statistically significantly lower valuations. On the contrary, we found having a degree to be insignificant. If one considers income to be a proxy for social class this was also insignificant in our results.

Table 7 shows the coefficients and *t*-statistics from an OLS regression to show the effect of background characteristics on TTO valuations. Experience of one's own ill health, being in work, having a degree, being a home owner, being a female in a household containing children (implying being a mother), and VAS for one's own current health, all increase the value assigned to a given health state. People who exhibit these characteristics are less willing to trade time for health improvements and hence implicitly assign higher values to those states.

Table 8 shows that people with children were slightly more concerned about being a burden on their family. Those in worse health found it easier to imagine a given hypothetical state, but surprisingly those who found it easier to imagine the states showed greater levels of inconsistency between VAS and TTO scores. Respondents who considered their ability to work in a given state also exhibited greater inconsistency between VAS and TTO.

6.4 Consistency across methods

Consistency was examined using the following approach: two-way correlations were derived, using each individual's responses, between the following variables: health state rankings, TTO score and VAS score. These were correlated using Spearman's rank correlation method. The mean Spearman correlation coefficient between ranking and TTO was -0.618 (the sign is negative, because low rank value corresponds to high valuation). The mean Spearman correlation coefficient between VAS and TTO was 0.582. These values suggest high levels of consistency between the three valuation methods. Note that since the states valued were generally very similar in severity, testing for consistency becomes more difficult. Comparison of TTO scores by interviewer found that one interviewer produced TTO scores that were statistically significantly a lot lower than the other two interviewers. It is necessary to have a 'balanced design' in which interviewers cover all health states evenly if we are to be able to detect such interviewer bias.

6.5 Time Preference

We attempted to assess time preference by asking respondents which of a number of different migraine scenarios they considered worst e.g. 26 days of migraine starting next month, 30 days of migraine in 15 years time or they are the same (testing 1% time preference). Questions were included to test 0%, 1%, 5% and 10% levels of time preference. Unfortunately 68 out of 108 respondents gave inconsistent responses to the four questions which made it impossible to calculate a time preference estimate for them. Of these inconsistent responders, 47 gave seemingly random responses; 16 stated that all scenarios were equal and five ranked them in reverse. Of the 40 consistent responders, the most prevalent level of time preference was >10% (12 responses).

7. Discussion and conclusions

The issues surrounding TTO valuation of swd may appear to be a highly specialised topic, of little practical importance. To the contrary, addressing these issues is central to the continued use of TTO. Nearly one third of the 243 health states described by the EQ-5D descriptive system have negative values in the MVH value set widely used in economic evaluation in the UK and beyond. Further, recall that the *mean* values for *most* health states are a product of values both positive and negative. The approach to valuation of swd is therefore highly relevant to the use of value sets as a defensible basis for public sector decision making.

Our results suggest that the lead time TTO is feasible for participants. The interviewers reported that the protocol was straightforward to administer. The values elicited using the lead time TTO had a high level of within-respondent consistency with the ranking and VAS evaluations of the states. One exception was the almost near universal ranking of dead as the worst state by our sample, in contrast to their subsequent TTO valuations. It seems possible that the ranking of dead as worst is a 'gut reaction' to what is the first task in the interview, and that this is modified once the participant engages with the task of imagining themselves to experience these states.

The re-scaled VAS valuations reported here are considerably higher than those reported in the MVH. This may be related to the set of states being considered in our study being exclusively very poor states of health, in contrast to the MVH study where these states were being evaluated alongside mild states. This potential violation of the independence of irrelevant alternatives has been noted in previous VAS studies (Ling-Hsiang and Kind 2008).

There is *prima facie* evidence that the use of a uniform procedure for sbd and swd overcomes discontinuities in values around 0. By employing a uniform procedure for states better and worse than dead, participants' responses can 'flip' from positive to negative values without the focusing effect created by the introduction of a separate valuation procedure. The characteristics of the distribution and variance of lead time TTO valuations requires further testing with a wider range of health states.

The effect of background characteristics on valuations suggests some quite complex influences. Those who have themselves experienced severe illness or have experience with poor health as a carer gave *lower* VAS values – whereas these same factors were associated with *higher* lead time TTO values, consistent with the theory that people with experience of ill health will have a greater understanding of adaptation and will value poor health states more highly. While the MVH study found social class and education to be the most significant influences on VAS valuations, we found having a degree and income (not included in the MVH study) were not related to VAS valuations. In addition to experience of one's own ill health, being in work, having a degree, being a home owner, being a female in a household containing children (implying being a mother), and VAS for one's own current health, all increased the lead time TTO values. This differs from the MVH study which found only age, sex and employment status to be significant. Some of these may be explained by the relatively high value placed on time (just 'being there') than on health. For example, maybe mothers, workers and those with mortgages to pay value their time more, and are therefore less willing to trade time for health improvements. This may mean that lead time TTO is more susceptible than conventional TTO to life stage concerns, where people trade off on the basis of when milestone events will happen (e.g. child reaching age 18; pay off mortgage), than on the basis of number of years of survival in different health states.

Some aspects of our findings need to be treated with caution. A considerable number of participants regarded the most severe states to be so bad that the 15 year lead time was insufficient to capture their anticipated disutility. Our results are based on the conservative[2] assumption that the minimum value in these instances was -4. However this was based on a simple 'yes/no' response about whether they would have been willing to trade if the lead time was extended to 20 *and* the duration of H_i was simultaneously reduced to 5. This was a pragmatic means of avoiding the complications involved in the use of alternative time scale fascias as in the pilot, and maintaining the overall 25 year profile participants had been considering in the TTO exercise. The -4 is not a product of iteration toward indifference and did not use any visual aid. However, even if it had, we would argue that altering the duration of H_i is *not* an appropriate means of coping with the issue of those who use up their lead time. As we noted in the Introduction in relation to the MVH and Torrance protocols, changing both the numerator and the denominator leads to extreme negative values. More fundamentally, a valuation procedure that relies on changing the duration of the state to be valued introduces the problem that the marginal value of time in poor health may not be constant, compounding the problem experienced by all TTOs that the marginal value of time in full health may not be constant.

The means of dealing with participants whose distaste for severe states is such that they exhaust the available lead time therefore requires further thought. Extending the lead time in such cases may be more readily handled using electronic props, avoiding the physical constraints of a TTO board.

Further, our attempts to measure time preference were not successful – the within-respondent inconsistencies suggest these questions were either too demanding or not worded sufficiently clearly. Therefore our results reported here do not control for time preference. This will be important to address in future work, as the introduction of lead time has the disadvantage that it pushes the state to be valued further into the future, potentially (depending on the durations involved) increasing the confounding effect of time preference on values.

Interpretation of the valuation data produced from the lead time TTO as being 'plausible' inevitably relies on expectations about the values for these states reported by other valuation studies, using different elicitation procedures. However, comparisons between mean lead time TTO values and those from other TTO studies, such as the MVH, is not straightforward. This entails either comparisons with valuation data that contain extreme negative numbers, or with artificially bounded ones, neither of which are valid comparisons. For instance, Figure 1 uses values re-scaled to -1 for the MVH study and figures bounded to -4 for our feasibility study in which we do not obtain numerical values when people prefer death to a 15 year lead time followed by poor health.

Finally, results of our initial experimentation with duration in the pilot suggests there are complex issues that remain to be explored regarding the interaction between duration and TTO valuation Although our feasibility study used a 10 year duration of the state to be valued, the rationale for selecting that (or indeed any single duration) is, admittedly, weak. More generally, there is also the issue that whether a state is better or worse than dead might itself be influenced by the time it is anticipated to be endured (Buckingham and Devlin 2008). This is an issue of considerable relevance to the use of TTO in economic evaluation.

[2] Conservative in the sense that -4 is the value associated with participants being indifferent between 20 years in full health and 5 in H_i, whereas their response could be consistent with greater disutility (see p.7). However, the -4 is *not* conservative as an estimate of the lead time TTO values that might apply – had we maintained a fixed denominator, but kept extending the lead time, the lead time would need to be 40 years at the point of indifference to yield this same value of -4.

Acknowledgments:

We are grateful to Anne Spencer, Economics Department, Queen Mary, University of London, and Angela Robinson, University of East Anglia, for their helpful suggestions and advice. We are also grateful to the administrative staff at City and Sheffield Universities and the members of the South Yorkshire community who participated in our study. This research was funded by a EuroQol Foundation grant. Ethical consent was obtained from the City University Senate Ethics Committee.

References

Buckingham K, Birdsall J, Douglas JG. (1996) Comparing three versions of the time trade-off: time for a Change? *Medical Decision Making*, Vol. 16, No. 4, 335-347.

Buckingham K (1993) A note on HYE (Healthy Years Equivalent). *Journal of Health Economics* 12(3): 301-9.

Buckingham K, Devlin N. (2008) *A note on the structure of utility space in time and health*. Discussion Paper 08/10, Economics Department, City University. www.city.ac.uk/economics/discussionpapers

Dolan P, Gudex C. (1995) Time preference, duration and health state valuations. *Health Economics* 4(4):289-99.

Dolan P. (1997) Modeling valuations for EuroQol health states. *Medical Care* 35(11):1095-108.

Lamers (2007) The transformation of utilities for states worse than death: consequences for the estimation of EQ-5D value sets. *Medical Care* 45(3) 238-244.

Lamers, L.M., McDonnell, J., Stalmeier, P.F.M., Krabbe, P.F.M., Van Busschbach, J (2006) The Dutch tariff: results and arguments for an effective design for national EQ-5D valuation studies. *Health Economics* 15(10):1121-1132.

Ling-Hsiang, C., Kind, P. (2008) *Ordinal or cardinal? The VAS strikes back*. Manuscript, www.york.ac.uk/inst/che/images/outcomes/Material_1.pdf

Patrick D, Starks H, Cain K, Uhlmann RF, Pearlman R. (1994) Measuring preferences for health states worse than death. *Medical Decision Making* 14: 9-18.

Robinson A, Spencer A. (2006) Exploring challenges to TTO utilities: valuing states worse than dead. *Health Economics* 15: 393-402.

Shaw JW, Johnson JA, Coons SJ. (2005) US valuation of the EQ-5D health states: development and testing of the D1 model. *Medical Care.* 43:203-220.

Stalmeier et al (2005) The gap effect: discontinuities in preferences around dead. *Health Economics* 14(7): 679-685.

Szende, A., Oppe, M. and Devlin, N. (2007) (Editors) *EQ-5D Value Sets: Inventory, Comparative Review and User Guide*, Springer Verlag.

Tilling C, Devlin N, Tsuchiya A, Buckingham K, (forthcoming), *Protocols for TTO Valuations of Health States Worse than Dead – A systematic review*, Health Economics and Decision Science Discussion Paper Series 08/09, University of Sheffield www.shef.ac.uk/scharr/sections/heds/discussion.html

Torrance GW. (1982) Multi-Attribute Utility Theory as a Method of Measuring Social Preferences for Health States in Long-Term Care. In *Values and Long Term Care*, pp. 127-56, R. Kane and R. Kane (eds.). Lexington Books Div., D.C. Health Publisher, Lexington, Mass.

Table 1. Characteristics of the feasibility study sample

		Feasibility study	MVH sample*
Sex:	Female Male	58% 42%	56% 44%
Age: average (standard deviation)	 *<44* *45-64* *65+*	46 (15.4) 45.0% 41.3% 13.7%	47 (18.4) 47.4% 28.9% 23.7%
Net Household income:	Under £10,000 £10,000 to £16,000 £16,001 to £20,000 £20,001 to £26,000 £26,001 to £40,000 £40,001 to £60,000 Over £60,000 No response	14.0% 8.4% 11.2% 4.7% 26.2% 12.2% 5.6% 17.8%	
Education after minimum school leaving age	No Yes	26.9% 73.1%	
Degree or equivalent professional qualification	No Yes	63% 37%	
Main activity	Housework Employment/self-employment Other Retired Seeking work Student	13.9% 55.5% 3.7% 18.5% 4.6% 3.7%	15.5% 49.0% 5.1% 22.5% 5.6% 2.3%
Home ownership status	Own home outright, or with mortgage Rent from local authority Rent from private sector	85.0% 8.4% 6.6%	
Experience of serious illness in you yourself	No Yes	80.6% 19.4%	68.3% 31.7%
Experienced serious illness in your family	No Yes	31.5% 68.5%	28.2% 71.8%
Experienced serious illness in caring for others	No Yes	57.4% 42.6%	72.4% 27.6%
Mean self-reported health on the EQ-VAS		0.81	0.86

Table 2. Summary of Ranking Results

EQ-5D state	Avge rank	% 1st	% 2nd	% 3rd	% 4th	% 5th	% 6th	% 7th	% 8th	% 9th	% 10th	% 11th	% 12th	Modal rank	Predicted % 1st rank
Dead	11.4	0	0	1	1	0	1	0	3	1	4	10	**78**	12	0.01
11111	1.2	**97**	0	1	0	0	1	0	0	0	0	1	0	1	82.60
13332	5.5	1	1	7	**32**	23	7	13	7	5	3	1	0	4	0.64
22222	3.6	1	5	**67**	13	7	1	0	2	1	0	2	0	3	2.63
32223	7.0	0	0	1	7	13	11	**27**	24	9	3	4	0	7	0.29
33323	8.8	0	1	3	3	1	4	4	11	27	**34**	11	0	10	0.09
11112	2.3	2	**87**	5	2	1	0	1	0	0	0	0	1	2	12.05
32211	5.5	1	1	12	25	11	**28**	9	4	2	3	3	0	6	0.66
33232	8.6	0	0	1	2	0	7	13	7	**40**	26	4	0	9	0.11
23232	5.5	0	2	7	24	**28**	16	8	7	3	3	1	1	5	0.62
33333	10.3	0	1	2	1	0	1	2	5	10	7	**53**	17	11	0.02
32232	7.1	0	0	1	7	11	14	**24**	29	4	8	2	0	7	0.28

NOTES:
- ‘Avge rank’ gives the average rank for each state, assuming rank scores are cardinal.
- ‘% 1st’ to ‘%12th’ represent the percentage of respondents giving a particular rank to a given state, and sum to 100% along the rows.
- Modal rank for a given state is indicated in bold, and presented under the ‘modal rank’ column.
- ‘Predicted % 1st rank’ gives the predicted probability of a given state being ranked first, based on the rank ordered logit regression results.
- n=92 except for states g and m where n=91.

Table 3. Raw and re-scaled VAS valuations from the feasibility study, compared with corresponding VAS valuations in the MVH

	Raw VAS, feasibility study			Raw VAS, MVH study			Rescaled VAS, feasibility study*			Rescaled VAS, MVH study*		
	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD
Dead	0	4.61	15.3	0	8.50	15.70	0.0	0.0	*n.a*	0.00	0.00	*n.a*
11111	100	96.29	13.9	100	98.70	4.80	1.0	1.0	*n.a*	1.00	1.00	*n.a*
13332	45	43.14	23.1	20	23.90	16.60	0.43	0.43	0.3	0.16	0.11	0.57
22222	70	66.05	20.2	50	52.30	17.30	0.69	0.66	0.24	0.50	0.45	0.37
32223	30	30.49	19.5	20	22.80	15.50	0.28	0.29	0.20	0.15	0.10	0.56
33323	15	19.79	19.1	10	13.90	12.40	0.10	0.17	0.22	0.07	-0.03	0.04
11112	90	84.18	18.1	87	82.40	15.20	0.86	0.86	0.21	0.87	0.81	0.23
32211	45	42.94	22.9	35	36.30	19.50	0.45	0.42	0.3	0.30	0.28	0.38
33232	20	21.43	17.4	14	16.20	12.70	0.16	0.19	0.2	0.10	0.01	0.71
23232	40	40.47	21.7	25	28.30	16.70	0.39	0.39	0.3	0.21	0.18	0.44
33333	9	13.68	19.1	2	5.60	9.10	0.1	0.11	0.2	0.00	-0.13	0.90
32232	30	30.35	20.1	20	23.40	15.90	0.29	0.29	0.3	0.17	0.06	0.77

* Rescaled values are anchored at 0 = dead and 11111 = 1, using U(Hi) = (Hi – d) / (11111- d). Four participants' data excluded where d > or = 11111

Table 4. TTO Values for Feasibility Study (baseline), Feasibility Study (minimum values restricted to -1.5) and MVH Study

	Lead Time Feasibility Study – Baseline Restricted to Minimum of -4.0			Feasibility Study with Values Restricted to Minimum of -1.5			Feasibility Study with Values Restricted to Minimum of -1.0 *					MVH Study: Transformed to Minimum of -1.0		
EQ-5D State	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD	LCL	UCL	Mean	Median	SD
13332	-0.2	-0.05	1.02	-0.11	-0.05	0.75	0.15	-0.01	0.47	0.06	0.24	-0.23	-0.38	0.55
22222	0.51	0.75	0.78	0.56	0.75	0.55	0.61	0.75	0.44	0.53	0.69	0.5	0.63	0.49
32223	-0.4	-0.15	1.17	-0.23	-0.15	0.74	0.05	-0.04	0.47	-0.04	0.14	-0.19	-0.28	0.56
33323	-0.82	-0.6	1.21	-0.59	-0.6	0.72	-0.13	-0.15	0.42	-0.21	-0.05	-0.39	-0.48	0.49
11112	0.79	0.95	0.36	0.79	0.95	0.36	0.8	0.95	0.31	0.74	0.86	0.82	0.93	0.29
32211	0.06	0.35	1.1	0.17	0.35	0.74	0.32	0.35	0.52	0.22	0.42	0.14	0.25	0.6
33232	-0.66	-0.55	1.28	-0.43	-0.55	0.78	-0.05	-0.14	0.47	-0.14	0.04	-0.33	-0.43	0.51
23232	0.05	0.23	0.81	0.07	0.23	0.72	0.26	0.23	0.46	0.17	0.35	-0.1	-0.08	0.59
33333	-1.16	-0.95	1.34	-0.79	-0.95	0.68	-0.25	-0.24	0.41	-0.33	-0.17	-0.54	-0.65	0.41
32232	-0.51	-0.47	1.11	-0.37	-0.47	0.76	0	-0.12	0.44	-0.08	0.08	-0.23	-0.38	0.57

* the feasibility study results were rescaled to achieve a minimum value of -1.0 by dividing all the negative scores by 4 (their maximum value).

Table 5. The effect of background characteristics on rankings

	COEFFICIENTS (t-values)									
Model	13332	22222	32223	33323	11112	32211	33232	23232	33333	control variable
baseline	-2.13	-3.35	-1.25	-0.24	-4.36	-2.13	-0.36	-2.1	-1.22	
	(-8.93)	(-13.56)	(-5.49)	(-1.07)	(-16.84)	(-8.95)	(-1.65)	(-8.85)	(-5.38)	
Controlling for own ill health	-2.13	-3.35	-1.25	-0.24	-4.37	-2.13	-0.36	-2.09	-1.22	-0.08
	(-8.93)	(-13.56)	(-5.48)	(-1.07)	(-16.84)	(-8.94)	(-1.65)	(-8.85)	(-5.38)	(-0.62)
Controlling for experience of ill health in others	-2.13	-3.35	-1.25	-0.24	-4.37	-2.13	-0.36	-2.1	-1.22	0.09
	(-8.93)	(-13.57)	(-5.49)	(-1.07)	(-16.85)	(-8.96)	(-1.65)	(-8.85)	(-5.38)	(0.87)
Controlling for experience as a carer	-2.13	-3.35	-1.25	-0.24	-4.36	-2.13	-0.36	-2.1	-1.22	0.02
	(-8.93)	(-13.56)	(-5.49)	(-1.07)	(-16.84)	(-8.95)	(-1.65)	(-8.85)	(-5.38)	(0.16)
Controlling for current smoking status	-2.13	-3.35	-1.25	-0.24	-4.37	-2.13	-0.36	-2.1	-1.22	0.07
	(-8.93)	(-13.56)	(-5.49)	(-1.07)	(-16.84)	(-8.96)	(-1.65)	(-8.85)	(-5.38)	(0.54)
Controlling for being in work	-2.13	-3.35	-1.25	-0.24	-4.36	-2.13	-0.36	-2.1	-1.22	-0.02
	(-8.93)	(-13.56)	(-5.48)	(-1.07)	(-16.84)	(-8.95)	(-1.65)	(-8.85)	(-5.38)	(-0.22)
Controlling for having a degree	-2.14	-3.36	-1.26	-0.24	-4.37	-2.14	-0.37	-2.1	-1.23	0.1
	(-8.96)	(-13.59)	(-5.51)	(-1.09)	(-16.86)	(-8.98)	(-1.67)	(-8.88)	(-5.4)	(0.98)
Controlling for being a home owner	-2.13	-3.35	-1.25	-0.24	-4.36	-2.13	-0.36	-2.09	-1.22	-0.04
	(-8.93)	(-13.56)	(-5.49)	(-1.07)	(-16.84)	(-8.94)	(-1.65)	(-8.84)	(-5.37)	(-0.31)
Controlling for being older than 45	-2.13	-3.35	-1.25	-0.23	-4.36	-2.13	-0.36	-2.09	-1.22	-0.06
	(-8.91)	(-13.55)	(-5.48)	(-1.06)	(-16.83)	(-8.94)	(-1.64)	(-8.84)	(-5.37)	(-0.64)
Controlling for being male	-2.13	-3.35	-1.25	-0.24	-4.37	-2.13	-0.36	-2.1	-1.22	0.08
	(-8.94)	(-13.56)	(-5.5)	(-1.07)	(-16.85)	(-8.95)	(-1.66)	(-8.86)	(-5.38)	(0.79)
Contolling for income over £26,000	-2.13	-3.35	-1.25	-0.24	-4.36	-2.13	-0.36	-2.09	-1.22	-0.07
	(-8.93)	(-13.56)	(-5.49)	(-1.07)	(-16.84)	(-8.94)	(-1.65)	(-8.84)	(-5.37)	(-0.72)
Controlling for being a 'mother'	-2.13	-3.35	-1.25	-0.24	-4.36	-2.13	-0.36	-2.1	-1.22	-0.03
	(-8.93)	(-13.56)	(-5.49)	(-1.07)	(-16.84)	(-8.95)	(-1.65)	(-8.85)	(-5.38)	(-0.28)
Controlling for visual analogue score for own health	0.01	-1.25	0.91	1.96	-2.29	0.01	1.83	0.04	0.94	0
	(0.02)	(-3.54)	(2.57)	(5.48)	(-6.52)	(0.03)	(5.14)	(0.13)	(2.66)	(0.47)

Table 6. The Effect of Background Characteristics on VAS values

		COEFFICIENTS (t-values)										
Model	constant	13332	22222	32223	33323	11112	32211	33232	23232	33333	control variable
Baseline	0.11	0.32	0.56	0.18	0.06	0.75	0.32	0.08	0.29	0.18	
	(4.35)	(9.45)	(16.28)	(5.31)	(1.75)	(22.03)	(9.28)	(2.41)	(8.42)	(5.29)	
Controlling for own ill health	0.3	0.32	0.56	0.18	0.06	0.75	0.32	0.08	0.29	0.18	-0.11
	(7.16)	(9.59)	(16.52)	(5.39)	(1.78)	(22.36)	(9.42)	(2.45)	(8.54)	(5.37)	(-5.64)
Controlling for experience of ill health in others	0.12	0.32	0.56	0.18	0.06	0.75	0.32	0.08	0.29	0.18	-0.01
	(3.73)	(9.45)	(16.28)	(5.31)	(1.75)	(22.03)	(9.28)	(2.41)	(8.42)	(5.29)	(-0.73)
Controlling for experience as a carer	0.16	0.32	0.56	0.18	0.06	0.75	0.32	0.08	0.29	0.18	-0.03
	(4.62)	(9.47)	(16.31)	(5.32)	(1.76)	(22.07)	(9.3)	(2.42)	(8.44)	(5.3)	(-2.17)
Controlling for current smoking status	0.1	0.32	0.56	0.18	0.06	0.75	0.32	0.08	0.29	0.18	0.02
	(1)	(9.45)	(16.29)	(5.31)	(1.75)	(22.04)	(9.28)	(2.41)	(8.42)	(5.29)	(1.13)
Controlling for being in work	0.12	0.32	0.56	0.18	0.06	0.75	0.32	0.08	0.29	0.18	-0.02
	(4.48)	(9.45)	(16.29)	(5.31)	(1.75)	(22.04)	(9.28)	(2.41)	(8.42)	(5.29)	(-1.14)
Controlling for having a degree	0.11	0.32	0.56	0.18	0.06	0.75	0.32	0.08	0.29	0.18	0
	(4.23)	(9.45)	(16.28)	(5.31)	(1.75)	(22.02)	(9.28)	(2.41)	(8.42)	(5.29)	(-0.02)
Controlling for being a home owner	0.16	0.32	0.56	0.18	0.06	0.75	0.32	0.08	0.29	0.18	-0.06
	(5.25)	(9.49)	(16.34)	(5.33)	(1.76)	(22.11)	(9.32)	(2.42)	(8.45)	(5.31)	(-2.92)
Controlling for being older than 45	0.11	0.32	0.56	0.18	0.06	0.75	0.32	0.08	0.29	0.18	-0.02
	(4.56)	(9.46)	(16.29)	(5.31)	(1.76)	(22.04)	(9.29)	(2.41)	(8.42)	(5.29)	(-1.37)
Controlling for being male	0.13	0.32	0.56	0.18	0.06	0.75	0.32	0.08	0.29	0.18	-0.06
	(5.15)	(9.51)	(16.38)	(5.34)	(1.77)	(22.17)	(9.34)	(2.43)	(8.47)	(5.32)	(-3.68)
Controlling for income	0.09	0.32	0.56	0.18	0.06	0.75	0.32	0.08	0.29	0.18	0.03
	(3.42)	(9.46)	(16.3)	(5.31)	(1.76)	(22.06)	(9.29)	(2.41)	(8.43)	(5.3)	(1.78)
Controlling for being a mother (a female in a household containing people < 18 years old)	0.09	0.32	0.56	0.18	0.06	0.75	0.32	0.08	0.29	0.18	0.06
	(3.45)	(9.51)	(16.38)	(5.34)	(1.76)	(22.17)	(9.34)	(2.43)	(8.47)	(5.32)	(3.66)
Controlling for own VAS	0.29	0.32	0.56	0.18	0.06	0.75	0.32	0.08	0.29	0.18	0
	(6.08)	(9.54)	(16.43)	(5.36)	(1.77)	(22.23)	(9.37)	(2.43)	(8.5)	(5.34)	(-4.47)

Table 7. The Effect of Background Characteristics on TTO Values

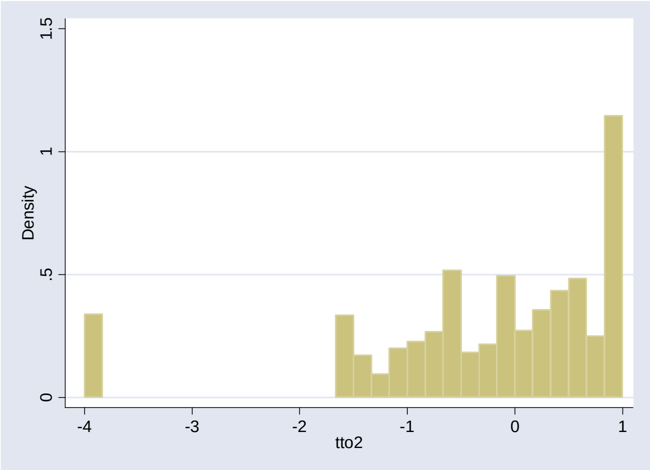
		COEFFICIENTS (t-values)										
Model	constant	13332	22222	32223	33323	11112	32211	33232	23232	33333	control variable
baseline	-1.16	0.97	1.68	0.77	0.35	1.95	1.22	0.51	1.21	0.65	
	(-11.38)	(6.72)	(11.61)	(5.29)	(2.41)	(13.45)	(8.48)	(3.52)	(8.39)	(4.52)	
baseline adjusted	-0.79	0.68	1.35	0.56	0.2	1.58	0.96	0.36	0.86	0.42	
	(-11.79)	(7.25)	(14.26)	(5.87)	(2.14)	(16.6)	(10.19)	(3.83)	(9.09)	(4.4)	
Controlling for own ill health	-1.77	0.97	1.68	0.76	0.35	1.95	1.22	0.51	1.21	0.65	0.34
	(-10.04)	(6.78)	(11.7)	(5.32)	(2.44)	(13.55)	(8.56)	(3.56)	(8.47)	(4.55)	(4.22)
Controlling for experience of ill health in others	-1.27	0.97	1.68	0.77	0.35	1.95	1.22	0.51	1.21	0.65	0.08
	(-9.23)	(6.72)	(11.61)	(5.29)	(2.41)	(13.46)	(8.49)	(3.52)	(8.39)	(4.52)	(1.12)
Controlling for experience as a carer	-1.03	0.96	1.68	0.77	0.35	1.95	1.22	0.5	1.21	0.65	-0.08
	(-7.1)	(6.71)	(11.61)	(5.29)	(2.4)	(13.45)	(8.48)	(3.51)	(8.39)	(4.51)	(-1.26)
Controlling for current smoking status	-1.15	0.97	1.68	0.77	0.35	1.95	1.22	0.51	1.21	0.65	-0.05
	(1)	(6.71)	(11.61)	(5.29)	(2.41)	(13.45)	(8.48)	(3.51)	(8.39)	(4.51)	(-0.63)
Controlling for being in work	-1.29	0.96	1.68	0.77	0.34	1.95	1.22	0.5	1.21	0.65	0.24
	(-12.05)	(6.74)	(11.68)	(5.32)	(2.41)	(13.5)	(8.52)	(3.52)	(8.42)	(4.52)	(3.75)
Controlling for having a degree	-1.26	0.97	1.68	0.77	0.35	1.95	1.22	0.51	1.21	0.65	0.27
	(-12.08)	(6.76)	(11.69)	(5.32)	(2.42)	(13.52)	(8.54)	(3.54)	(8.46)	(4.54)	(4.01)
Controlling for being a home owner	-1.48	0.96	1.67	0.76	0.35	1.95	1.22	0.5	1.21	0.65	0.37
	(-11.74)	(6.76)	(11.68)	(5.29)	(2.42)	(13.52)	(8.54)	(3.54)	(8.45)	(4.55)	(4.21)
Controlling for being older than 45	-1.21	0.97	1.68	0.77	0.35	1.95	1.22	0.51	1.21	0.65	0.1
	(-11.34)	(6.72)	(11.63)	(5.3)	(2.42)	(13.46)	(8.49)	(3.52)	(8.4)	(4.53)	(1.5)
Controlling for being male	-1.12	0.97	1.68	0.77	0.35	1.95	1.22	0.51	1.21	0.65	-0.11
	(-10.61)	(6.72)	(11.63)	(5.31)	(2.41)	(13.48)	(8.49)	(3.52)	(8.4)	(4.52)	(-1.62)
Controlling for income	-1.23	0.97	1.68	0.76	0.35	1.95	1.22	0.51	1.21	0.65	0.1
	(-11.17)	(6.72)	(11.61)	(5.28)	(2.41)	(13.45)	(8.49)	(3.52)	(8.4)	(4.52)	(1.55)
Controlling for being a mother (a female in a household containing people < 18 years old)	-1.21	0.97	1.68	0.77	0.35	1.95	1.22	0.51	1.21	0.65	0.14
	(-11.58)	(6.72)	(11.64)	(5.31)	(2.41)	(13.48)	(8.49)	(3.52)	(8.41)	(4.52)	(2.05)

Table 8. Statistically significant interactions between background characteristics, participant feedback and valuation data in the feasibility study

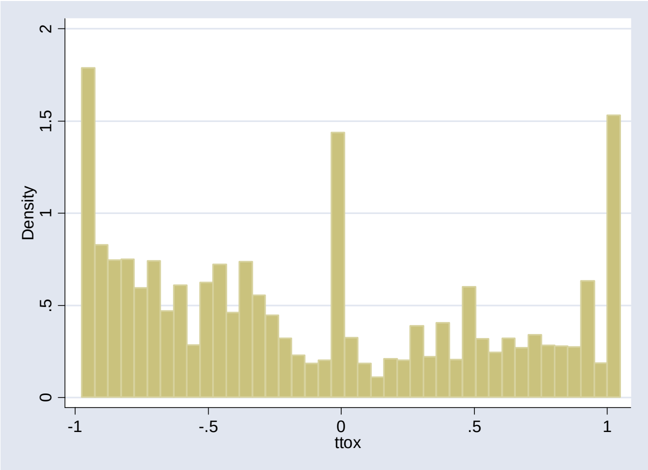
Variable 1	Variable 2	X^2 Test	Relationship
Number of children in household[3]	Being a burden on my family was an important consideration[4]	0.072	More important for those with 2 or more children but still important for majority of people with no children
VAS of own health[5]	I found it difficult to imagine the states[2]	0.066	Those in worse health found it easier to imagine the states
Consistency between VAS and TTO scores (sprho)[6]	I found it difficult to imagine the states[2]	0.002	Those who found it difficult to imagine the states had more consistent TTO and VAS valuations for each state.
Consistency between VAS and TTO scores (sprho)[4]	Ability to work an important consideration[2]	0.011	VAS and TTO valuations for each state were more consistent amongst those who did not consider ability to work important

Figure 1. Distribution of TTO results

(a) Current study, all states, untransformed values



(b) MVH study, states used in current study, transformed to -1



[3] Variable coded: 0 children, 1, 2 or more
[4] The responses to this question ranged from Agree (1) to Disagree (5). For the cross tab the variable was coded dichotomously: Agree (1-2), Disagree (3-5).
[5] Variable coded: <71 (0), 71-80 (1), 81-90 (2), 91-100 (3)
[6] Variable coded: Consistent (>0.8), Inconsistent (<0.8)