New findings from the TTO for income approach to elicit willingness to pay for a QALY

Attema, Arthur and Krol, Marieke and van Exel, Job and Brouwer, Werner

Erasmus University Rotterdam

17 December 2014

New findings from the TTO for income approach to elicit willingness to pay for a QALY

Arthur E. Attema[a], Marieke Krol[b], Job van Exel[b], Werner B.F. Brouwer[b]

[a] (Corresponding author) iBMG, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands. E-mail: attema@bmg.eur.nl, --31-10.408.91.29 (O); --31-10.408.90.81 (F)
[b] iBMG, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, the Netherlands.

December, 2014

ABSTRACT. Willingness to pay (WTP) elicitations suffer from various methodological problems. This paper tests a recently proposed alternative approach to value WTP for health, making use of trade-offs between income and lifetime or quality of life. We apply three experimental elicitation procedures and analyze the responses under an additive and a multiplicative utility function over health and income. We report several interesting results. First, the data are highly skewed, but if we trim the 5% lowest and highest values, we obtain plausible WTP estimates. Second, the results differ considerably between procedures, indicating that WTP estimates are sensitive to the assumed utility function. Third, respondents appear to be loss averse for both health and money, which is consistent with assumptions from prospect theory. Finally, our results also indicate that respondents are more willing to trade quality of life than life years.

Key Words: loss aversion, time tradeoff method, QALY, utility of health and wealth, willingness to pay

JEL CLASSIFICATION: B41, D03, I10

1. Introduction

Economic evaluations provide information on costs and effects of health technologies. Within economic evaluations, health effects are typically expressed in Quality Adjusted Life Years (QALYs). The QALY is a uniform outcome measure of health benefit that combines length of life with quality of life (QoL). By expressing health outcomes with a uniform measure, outcomes can be compared across different diseases and treatments, which can be helpful for decision makers in the process of making reimbursement decisions.

While operating under budgetary constraints and pressure, advisory bodies, such as the National Institute for Health and Clinical Excellence (NICE, 2008) in England and the National Health Care Institute (College Voor Zorgverzekeringen (Dutch board of insurers), 2006) in the Netherlands, are in search for the shadow price of a QALY (Culyer et al., 2007). However, these two bodies use different shadow prices: NICE bases its shadow price upon forgone health (Claxton et al., 2010), whereas the Dutch National Health Care Institute bases it upon the consumption value of health (College Voor Zorgverzekeringen (Dutch board of insurers), 2006).

In the first case, the value of health is determined by comparing the expected health gains of a health intervention to the health that is likely forgone elsewhere due to the displacement of activities within a fixed budget. This approach is also labelled as adopting a health care perspective, focussing only (or primarily) on costs to the health care sector and the health effects of an intervention. Cost-effectiveness analyses may suffice to prioritize healthcare in this case, operating under an exogenous budget constraint that is imposed by a higher authority (Claxton et al., 2010; 2011). In general, the decision rule then indicates that only when the health gained exceeds the health forgone, a new intervention should be adopted. It is not possible within this framework to judge whether the budget itself has been set appropriately.

In the latter case, the value of health is determined by assessing the amount of consumption that individuals are willing to give up to improve health (Claxton et al., 2010). This approach may be related to adopting a societal perspective in performing economic evaluations, taking into account the broader societal costs and benefits of health interventions. Countries considering using this decision framework require a monetary estimate of the (consumption) value of health. The decision rule then becomes that the monetary value of the health (welfare gained) produced should exceed the monetary value of the costs (welfare sacrificed). As long as this rule is followed in adopting and applying technologies, the appropriate budget follows from these decisions. In this paper we focus on the estimation of the consumption value of health and, hence, we seek to estimate the monetary value of a QALY (MVQ).

Two kinds of willingness to pay (WTP) approaches have frequently been used to estimate the MVQ. The first approach has been to elicit the WTP for a reduction in the risk of death and then calculate the value of a life, from which the MVQ can be inferred (Abelson, 2003; Hirth et al., 2000; Johannesson and Meltzer, 1998; Mason et al., 2009). The second approach has been to elicit the WTP for changes in health status directly (Bobinac et al., 2010; 2014; Gyrd-Hansen, 2003; Johannesson and Johansson, 1997; Johnson et al., 1998; King et al., 2005; Lundberg et al., 1999; Pennington et al., 2013; Pinto-Prades et al., 2009; Robinson et al., 2013).

However, the method of WTP has several known problems, including: insensitivity to scope (Bobinac et al., 2012; Olsen et al., 2004); strategic behaviour (Hackl and Pruckner, 2005); the restriction of personal income or 'ability to pay' (O'Brien and Drummond, 1994); protest responses (Dalmau-Matarrodona, 2001); and dependence on the elicitation method (Frew et al., 2003), the payment vehicle (Hayes et al., 1992) as well as the order of the questions if more than one outcome is being valued (Stewart et al., 2002). Another shortcoming of most WTP elicitations is their inability to take account of reference-dependency, which has often been demonstrated to play a considerable role in people's decisions and valuations (Kahneman and Tversky, 1979; Moffett and Suarez-Almazor, 2005; Starmer, 2000; Treadwell and Lenert, 1999; van Exel et al., 2006).

Recently, Tilling et al. (2014) suggested an alternative approach to estimate the MVQ, based upon a Time Trade Off (TTO) exercise. In that method, people are asked to choose between living longer (in some fixed health state) with less income and living shorter (in that same health state) with more income. Thus, a trade-off is made between length of life (in a particular health state) and income, which allows investigation of the implicit monetary value given to QALYs. Tilling et al. (2014) performed a first test of feasibility of this new approach and found that it may be possible to generate satisfactory WTP estimates, but they experienced a number of drawbacks. One of them was the need to specify a utility function over income, length of life and QoL. Tilling et al. (2014) estimated WTP assuming an additive lifetime utility function, which may be too restrictive (Domeij and Johannesson, 2006; Finkelstein et al., 2013; Sloan et al., 1998; Viscusi and Evans, 1990). Therefore, in this paper we investigate this WTP-TTO approach more extensively in a representative sample of the Dutch population, using a multiplicative utility function in the computation of WTP and allowing for reference-dependence and loss aversion.

2. Methods

Tilling et al. (2014) assumed an additive function W(.) over healthy life years (H) and income (Y):

$$W(H,Y) = U(H) + Y \tag{1}$$

That is, individuals derive value from their lifetime and have a linear utility function over income. This specification was used earlier by Eeckhoudt et al. (1998). The advantage of this function is that it becomes straightforward to elicit a monetary value of the utility of perfect health. The pitfall is that it is descriptively less accurate. In particular, assuming this utility function implies independence of consumption utility from the level of health, which was one of the 'impossibility theorem criteria' set out by Dolan and Edlin (2002). Moreover, the empirical literature tends to reject this assumption in favour of a multiplicative utility function over health and income. Indeed, there is evidence that marginal utility of wealth increases with health and longevity, which is impossible under an additive function (Domeij and Johannesson, 2006; Finkelstein et al., 2013; Sloan et al., 1998; Viscusi and Evans, 1990).[1] Hence, we also study the following utility function over health and income:

$$W(H,Y) = U(H) \times V(Y) \tag{2}$$

Bleichrodt and Quiggin (1999) have given the axiomatic foundations for this function. The simplest configuration would be to take both U(H) and V(Y) to be linear, but this is not very realistic. It is more likely that marginal utility decreases with income, i.e., $V'(Y)<0$. Here, we model this by considering a power utility function $V(Y) = Y^{\alpha}$, with α as a measure of the utility curvature of income and $V(Y) = \ln(Y)$ for α=0. Therefore, our lifetime utility function will take the form:

[1] Tengstam (2014) instead found evidence that marginal utility of income is decreasing with health.

$$W(H,Y) = H \times Y^{\alpha} \quad (3)$$

Empirical support for this function was provided by Levy and Nir (2012). We take into account the possibility of health being less than perfect, by assuming the QALY model: H=Q×T, with T as the number of life years, and Q as the QoL on a cardinal scale with 0 indicating a health state as bad as death and 10 indicating full health.

2.1.TTO for income – classical approach

The TTO for income method lets a subject compare a particular remaining lifetime T with some income level Y_A to another amount of remaining lifetime X with another income level Y_B. A possible scenario would be to assume T=10 years of life with QoL Q and annual income Y_1 and to ask for the amount of remaining lifetime X which would render indifference in case the income level would increase to a higher level Y_B, while QoL remains stable at Q (*TTO1*). Under the multiplicative model (Eq. 3), this would result in the following equality:

$$10 \times Q \times {Y_{\mathrm{A}}}^{\alpha} = X \times Q \times {Y_{\mathrm{B}}}^{\alpha} \quad (4)$$

From this, we can compute an estimate of α:

$$\left(\frac{Y_{\mathrm{A}}}{Y_{\mathrm{B}}}\right)^{\alpha} = \frac{X}{10} \Leftrightarrow \alpha = \frac{\ln(10)-\ln(X)}{\ln(Y_{\mathrm{B}})-\ln(Y_{\mathrm{A}})} \quad (5)$$

with α>0. Having this estimate, we can continue to infer an estimate of the WTP for one year in full health (*WTP[YFH]*). For example, in case living 9 years with the higher income Y_B would give equal lifetime utility as the initial scenario with 10 years and income Y_A, both in full health (i.e., Q=10), the estimated value of α would be:

$$10 \times 10 \times {Y_{\mathrm{A}}}^{\alpha} = 9 \times 10 \times {Y_{\mathrm{B}}}^{\alpha} \Leftrightarrow Y_{\mathrm{B}} = \left(\frac{10}{9}\right)^{1/\alpha} \times Y_{\mathrm{A}} \quad (6)$$

WTP for a healthy life year is then given by the additional lifetime income people demand in return for reducing life by one year, corrected for their QoL:

$$WTP(YFH) = \frac{(9Y_9 - 10Y_A)}{Q} =$$

$$\frac{\left(9\left(\frac{10}{9}\right)^{1/\alpha} - 10\right)Y_A}{Q} \quad (7)$$

Alternatively, under the additive utility function (Eq. 1), the indifference above (with X=9 and Q=10) will be evaluated by:

$$10 \times 10 \times WTP(YFH) + 10 \times Y_{\mathrm{A}} = 9 \times 10 \times WTP(YFH) + 9 \times Y_{\mathrm{B}} \quad (8)$$

Solving Eq. 8 for WTP(YFH) yields:

$$WTP(YFH) = \frac{9 \times Y_{\mathrm{A}} - 10 \times Y_{\mathrm{B}}}{(10-9) \times Q} = \frac{9 \times Y_{\mathrm{A}} - 10 \times Y_{\mathrm{B}}}{Q} \quad (9)$$

2.2. TTO for income – behavioural economic approach

A large body of evidence has emerged suggesting that people deviate from several rationality assumptions underlying neoclassical economic theory. One such deviation is that individuals tend to behave according to prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992; van Exel et al., 2006). In particular, they often form reference points and handle gains and losses as seen from this reference point differently. There is evidence that this behaviour also occurs in health-related decision making (Attema et al., 2013; Bleichrodt and Pinto, 2002; Bleichrodt et al., 2003). In order to accommodate this possibility, we analyse our data under this assumption from prospect theory as well.

Preferences become reference-dependent if we assume prospect theory, which requires separate formulations for gains and losses. In particular, we investigate reference-dependency by the model proposed by Shalev (2002), which for income culminates into:

$$U(Y) = \begin{cases} u(Y) & \text{if } Y \geq Y_0 \\ u(Y - \lambda_M(Y_0 - Y)) & \text{if } Y < Y_0 \end{cases} \quad (10)$$

With λ_M a loss aversion index for monetary outcomes and Y_0 the status quo. Although the utility function may be different for gains and losses, e.g. u(Y) = Y^α for gains and u(Y) = $-(-Y)^\beta$ losses, with α,β>0, for simplicity we assume they are the same. Extending this model to health yields:

$$U(H) = \begin{cases} u(H) & \text{if } H \geq H_0 \\ u(H - \lambda_H(H_0 - H)) & \text{if } H < H_0 \end{cases} \tag{11}$$

with λ_H a loss aversion index for health outcomes.

In the next three sections (2.2.1 – 2.2.3) we describe the three experimental procedures that will be applied in this study and two hypotheses to be tested based on these procedures.

2.2.1. TTO income gain and health loss

Suppose we apply TTO1 with T=10 years, Q=the respondent's own current health state, Y_A=C and Y_B=L (with L>C), and we ask for the number of years X_1 with income L rendering indifference. According to prospect theory, respondents then have to trade off a gain in income against a loss in lifetime. If we assume {T Years, Y_A} to be the reference point, this involves comparing the status quo against a mixed prospect, which would be evaluated by:

$$10 \times Q \times C^\alpha = [X_1 - \lambda_H(10 - X_1)] \times Q \times L^\alpha \tag{12}$$

Solving this expression for X_1 gives:

$$X_1 = \frac{10}{1+\lambda_H}\left[\lambda_H + \left(\frac{C}{L}\right)^\alpha\right], \tag{13}$$

which is increasing in λ_H. Therefore X_1 will be higher for people who are loss averse (λ_H>0) than for people who are loss neutral (λ_H=0). In the classical approach described in section 2.1 loss aversion is ignored implicitly assuming λ_H=0. Consequently, the effect of loss aversion will be picked up by our estimate of α (Eq. 5), which is decreasing in X_1 and, hence, will be lower if people are loss averse than if they are not. As derived in the Appendix, the real estimate of α is given by $\alpha = \frac{ln\left(\frac{X_1(1+\lambda_H)}{10}+\lambda_H\right)}{\ln(C)-\ln(L)}$, which requires knowledge of λ_H. Therefore, because our estimated α is decreasing in X_1, and X_1 increases with λ_H, the classical approach can be expected to generate an underestimation of the true α in case of loss aversion and, hence, an overestimation of WTP for a QALY (Eq. 7).

In case of the additive model, reference-dependence gives the following evaluation:

$$10 \times Q \times WTP(YFH) + 10 \times C = [X_1 - \lambda_H(10 - X_1)] \times Q \times WTP(YFH) + X_1 \times L \quad (14)$$

Solving for X_1 gives:

$$X_1 = \frac{10\times[(Q\times WTP(YFH))(1+\lambda_H)+C]}{(Q\times WTP(YFH))(1+\lambda_H)+L}, \quad (15)$$

which is again increasing in λ_H. Because WTP is increasing in X_1 (Eq. 9), we again predict an overestimation of WTP for a QALY in case of loss aversion.

2.2.2. *TTO income loss and health loss*

Now suppose we apply TTO1 with T=10 years and Y_B=L (with L>C), as in 2.2.1, but Y_A=S (with S<C), and ask for X_2 which would render indifference in case the income level would increase to the present level C (*TTO2*). If we assume {T Years, Y_A} is still the reference point, the first option now entails a loss in income, whereas the second option still entails a loss in health. In other words, we are now comparing a loss in the monetary domain to a loss in the health domain. Indifference between the two options can then be evaluated by:

$$10 \times Q \times [S - \lambda_M(C - S)]^\alpha = [X_2 - \lambda_H(10 - X_2)] \times Q \times C^\alpha, \quad (16)$$

which gives a different solution for X_2 than we had for X_1 in the first procedure (Eq. 15):

$$X_2 = 10\left[\frac{(C+\lambda_M(S-C))^\alpha - C^\alpha}{\lambda_H C^\alpha} + 1\right] \quad (17)$$

X_2 is increasing in λ_H again, but at the same time decreasing in λ_M. In other words, the two loss aversion coefficients are opposing forces in determining X_2 and the qualitative effect of loss aversion on X_2 will therefore depend on the relative values of λ_H and λ_M. Consequently, the estimate of α (Eq. 5) is expected to be smaller in TTO2 than in TTO1. Since α is predicted to be an underestimation in TTO1, this underestimation would then be even smaller in TTO2, perhaps changing into an overestimation if λ_H is high enough. We therefore formulate the following hypothesis:

Hypothesis 1 (TTO1 vs. TTO2): $\alpha_1<\alpha_2$.

Finally, incorporating sign-dependence into the additive model gives the beneath expression for X_2, which is increasing in λ_H and decreasing in λ_M, yielding the same predictions as for the multiplicative model:

$$X_2 = \frac{10\times[(\text{Q}\times\text{WTP(YFH)})(1+\lambda_H)+\text{S}(1+\lambda_M)-\lambda_M\text{C}]}{(\text{Q}\times\text{WTP(YFH)})(1+\lambda_H)+\text{C}}. \quad (18)$$

2.2.3. QTO

A third possibility to elicit MVQ is a new variation to the common TTO for income procedure: the Quality Trade Off (QTO). This procedure varies QoL instead of life duration. Suppose we apply QTO with T=10 years in full health, Y_A=C and Y_B=L (with L>C), and we ask for the QoL score X_3 with T=10 years in income L rendering indifference. Health status is described on a 10-point scale, with 10 representing perfect health and 0 a health state as bad as death. If there is no reference-dependency, this indifference can again be evaluated by Eq.3, yielding:

$$10 \times 10 \times WTP(YFH) \times C^{\alpha} = 10 \times X_3 \times WTP(YFH) \times L^{\alpha} \quad (19)$$

$$\alpha = \frac{\ln(10)-\ln(X_3)}{\ln(L)-\ln(C)} \quad (20)$$

Because according to the QALY model T and Q are fully exchangeable, meaning that living 10 years with QoL 9 is equivalent to living 9 years with QoL 10, solving Eq. 19 for WTP(YFH) yields the same result as Eq. 7. Therefore, this model predicts WTP and α to be the same in TTO1 and QTO. In other words X_1 is predicted to be equal to X_3.

If prospect theory holds instead, respondents have to trade off a gain in income against a loss in QoL. Assuming {10 years in full health, C} to be the reference point, this again involves comparing the status quo against a mixed prospect, which under the multiplicative model would be represented by:

$$10 \times 10 \times C^{\alpha} = 10 \times 10 \times \left[X_3 - \lambda_Q(1 - X_3)\right] \times L^{\alpha} \tag{21}$$

This expression can be solved for X_3:

$$X_3 = \frac{10}{1+\lambda_Q}\left[\lambda_Q + \left(\frac{C}{L}\right)^{\alpha}\right], \tag{22}$$

Comparing (22) to (13), it becomes evident that X_1 and X_3 are expected to differ only to the extent that loss aversion for QoL differs from loss aversion for life duration.

Under the additive model, reference-dependence gives:

$$10 \times 10 \times WTP(YFH) + 10 \times C = 10 \times \left[X_3 - \lambda_Q(1 - X_3)\right] \times WTP(YFH) + 10 \times L \tag{23}$$

Solving Eq. 23 for X_3 gives:

$$X_3 = 10 \times \left(1 - \frac{L-C}{(WTP(YFH))(1+\lambda_Q)}\right), \tag{24}$$

which is again increasing in λ_Q.

It is not clear beforehand whether loss aversion is stronger for life duration or for QoL. Intuitively, people may be more reluctant to give up lifetime, which would translate into more loss aversion for life duration than for QoL, but no firm evidence is available on this point. Consequently, our second hypothesis is the following:

Hypothesis 2 (TTO1 vs. QTO): $\boldsymbol{\lambda_H=\lambda_Q}$. We will test this hypothesis by comparing X_1 and X_3, which are predicted to be equal if $\lambda_H=\lambda_Q$, as derived earlier.

3. Experiment

3.1. Subjects and income levels

A total of 550 subjects representative for the Dutch adult population in terms of gender, age and level of education participated in the experiment. The study presented here was part of a larger experiment that included two other studies.

Before the experiment started, subjects were among others asked for: their current net household income (called C hereafter), the net monthly income that would be sufficient to just make ends meet while staying in their current house (*subsistence income*, called S hereafter), and the net monthly income they would need to be able to live a comfortable life without any worries (*luxury income*, called L hereafter).

3.2. Stimuli

3.2.1. TTO1

In TTO1, respondents were asked to choose to live T=10 more years in their current health state and their current monthly salary, multiplied by 12 to get yearly income Y_A=C, or to live an amount $X_1 \leq 10$ years in their current health state Q (as measured by a visual analogue scale in the beginning of the experiment) but with a higher income (Y_B=L).

TTO1: Trading years to achieve an income gain in current health
Suppose you can choose between the following two options.
Option A.
*"You live for 10 years in your current health state with a net monthly income of [*C/12*], without any changes to it. Then you die."*
Option B.
*"You live for X years in your current health state with a net monthly income of [*L/12*], without any changes to it. Then you die."*

Hence, TTO1 elicited the number of life years X_1such that the subject would be indifferent between (10 years, C) and (X_1 years, L), which gives the estimates of α and WTP according to Eq.5 and Eq.7 under the multiplicative model and according to Eq.9 under the additive model with X= X_1.

3.2.2. TTO2

TTO2 was as explained at the end of the previous section, Y_A= S.

TTO 2: Trading years to achieve an income gain in current health

Suppose you can choose between the following two options.

Option A.

*"You live for 10 years in your current health state with a net monthly income of [*S/12*], without any changes to it. Then you die."*

Option B.

*"You live for X years in your current health state with a net monthly income of [*C/12*], without any changes to it. Then you die."*

TTO2 gives the estimates of α and WTP for the multiplicative [additive] model as provided in Eqs. 5 and 7 [9] again, with Y_A=S, Y_B=C, and X=X_2.

3.2.3. QTO

For income we again used current income C and luxury income L, whereas we used 10 years of life in both options.

QTO: Trading quality to achieve an income gain during 10 remaining years.

Suppose you can choose between the following two options.

Option A.

*"You live for 10 years in perfect health state (10 on a scale of 0 to 10) with a net monthly income of [*C/12*], without any changes to it. Then you die."*

Option B.

*"You live for 10 years in moderate health (X on a scale of 0 to 10) with a net monthly income of [*L/12*], without any changes to it. Then you die."*

3.3. Procedure

In the first choice, X was always equal to 10 years (life duration part)/QoL points (QoL part). Because monotonicity implies dominance of Option B in this situation, we would expect respondents to opt for B here. In case one chose A, we asked whether they really preferred 10 years with C to 10 years with L. If so, they received the next question and a missing value was stored for them. Otherwise, they received the original question anew. If respondents were indifferent, a value of 10 was saved. If B was chosen, X was randomly lowered to 3, 5, or 7 years/QoL points. The respondent could then choose A or B again or express indifference. In case of indifference, the provided value of X was the elicited indifference point. If A or B was chosen, the respondent had to indicate the value of X such that A and B were equally attractive to them by using a scroll bar, where the range of the scroll bar was censored by the previous choice. For example, if the respondent received X=3 in the second choice and then opted for A, the scroll bar was censored between 3 and 10, whereas it was between 0 and 3 if they opted for B.

The experiment was conducted by a professional internet sampling company (Survey Sampling International). This company has much experience with internet surveys and a large representative database of subjects. The subjects were rewarded with a small monetary amount to be given to a charity fund of their choice, upon completion of the questionnaire.

Income was measured on a categorical scale (with "€999 or less" as lowest category, "€8,000 or more" as highest category, and eleven €500 intervals in-between). We used the midpoint of the chosen scale as the amount to be used in TTO questions. Whenever someone expressed subsistence income to be above current income, or luxury income below current income, we replaced these values in the TTO questions in order to enable sensible trade-offs. In particular, S was replaced by half of current income and L was replaced by twice the amount of current income.

3.4. Analysis

As pointed out by Gyrd-Hansen and Kjær (2012), there tends to be a lot of heterogeneity in WTP for a QALY estimates. They demonstrate that, because of this heterogeneity, the choice of the analytical approach can make for a large difference in WTP estimates. In particular, they summed the individual WTP estimates and divided this amount by the sum of the considered QALY gains (aggregated approach or 'ratio of means'). They compared this approach to the disaggregated approach ('mean of ratios'), where they divided the WTP by the associated QALY gain for each individual separately, and observed large differences. In the disaggregated approach, it was not possible to include non-traders, because their QALY gain was zero. In our case, the aggregated approach means a division of the sum of the income differences by the sum of the life time reductions. The disaggregated approach instead implies a division of the income difference by the reduction of life time for each respondent. These approaches are also likely to generate different results, especially because we have a lot of non-traders, who could be included in the aggregated approach, but not in the disaggregated approach.

We compared the individual WTPs in the disaggregated approach by means of a nonparametric Wilcoxon signed ranks test (a Kolmogorov-Smirnov test indicated that the WTP estimates were not normally distributed, $p<0.01$). For TTO1 [TTO2, QTO], 247 [180, 148] non-traders had to be excluded from this analysis. Two respondents who had a subsistence income level of S=0 were also excluded, because α could not be computed for them. Finally, respondents who traded off all available life years had to be excluded since these people would have no life time left (1 in TTO1, 8 in TTO2, and 2 in QTO).

4. Results

Table 1 presents some demographic variables of our sample. These numbers indicate representativeness for the Dutch adult population according to age, gender and education.

Table 1. Summary statistics

Variable	Percentage	Mean	SD	Minimum	Maximum
Age		45.6	15.02	18	75
Gender (% male)	49.3				
Children (%yes)	57.5				
Number of children (among people with children, n=316)		2.22		1	21
Income groups:					
<€1000	14.0				
€1000-<€2000	37.1				
€2000-<€3000	28.9				
€3000-<€4000	13.3				
>€3999	6.7				
Education:					
Lower	28.6				
Middle	41.6				
Higher	29.8				
Health status					
EQ-5D (Dutch tariff)		0.82	0.21	-0.329	1
VAS		76.75	17.75	9	100
Completion time (mins.)		16.1	6.2	5	44.9

Table 2 classifies the respondents in different groups: non-traders, over-traders (causing a negative WTP) and respondents with zero WTP. Row F shows the net sample size for each method. Rows G and H display the mean number of years [QoL] given up in TTO1 and TTO2 [QTO]. These numbers indicate that respondents sacrifice more years in TTO2 than in TTO1, and that the relative sacrifice is similar for TTO2 and QTO. Furthermore, a high number of non-traders appeared in all three tasks. These respondents may be viewed as people who “are not willing to play the game”.

Table 2. Overview

		WTP1 (L-C)	WTP2 (C-S)	WTP3 (L-C QoL)
A	Non-traders	247	180	148
B	Negative WTP	77	111	0 (add) 151 (mul)
C	Subjects trading off all years/quality	1	8	2
D	Zero WTP	12	59	0
E	S=0		2	
F	Net sample size (550-A-C)	302	360	400
G	Mean years/QoL given up (incl. non-traders)	1.96	2.81	28.6%
H	Mean years/QoL given up (excl. non-traders)	3.56	4.21	39.2%

Table 3 presents the WTP estimates under the assumption of the additive model. The observation of Table 2 of more life years given up to move from a subsistence income to their current income, than to move from their current income to a luxury income, clearly translates into a lower WTP estimate in the former task than the latter. In addition, the substantial number of over-traders results in a negative WTP in the additive model, explaining the low median WTP.

Table 3. WTP estimates additive model

	WTP1 (L-C)	WTP2 (–C-S)	WTP3 (L-C QoL)
Mean (disaggregated approach)	234,465	55,641	132,322
Median (disaggregated approach)	20,563	3,542	42,000
WTP using means (incl. non-traders) (aggregated approach)	117,611	16,916	98,708
Mean trimmed data (5% upper and 5% lower) (disaggregated approach)	78,629	13,377	77,114
Median trimmed data (disaggregated approach)	20,563	3,875	42,000
WTP using means trimmed data (aggregated approach)	86,518	17,833	71,493
WTP using medians trimmed data (aggregated approach)	401,250	5,000	62,069

Table 4 gives the estimates obtained under the multiplicative model. This table shows a similar pattern across methods, but at the same time the estimates differ substantially from those of the additive model. In particular, the means explode when estimating the multiplicative model, due to some outliers. These are less influential in the aggregated approach, giving much more conservative estimates there. Moreover, the medians are less vulnerable to outliers than the means; they are even lower when fitting the multiplicative model than when fitting the additive model. In order to remove the inflating effect of the outliers, we also analyzed the data using a trimmed dataset, where we removed the 5% highest and 5% lowest WTP ratios. In general, both models generate considerably higher means than medians, indicative of a high degree of skewness, which is common in WTP studies (Smith, 2001; Stewart et al., 2002).

Table 4. WTP estimates multiplicative model

	α_1	WTP1 (L-C)	α_2	WTP2 (C-S)	α_3	WTP3 (L-C QoL)
Mean (disaggregated approach)	1.01	7,82e14	1.12	2,931,121	1.25	1,985,200
Median (disaggregated approach)	0.62	14,969	0.86	2,604	0.79	4,060
WTP using means (aggregated approach)	0.29	96,503	0.60	14,072	0.46	34,564
Mean trimmed data (5% upper and 5% lower) (disaggregated approach)		87,328		11,596		28,675
Median trimmed data (disaggregated approach)		14,969		2,254		4,061
WTP using means trimmed data (aggregated approach)	0.33	75,288	0.57	14,807	0.55	22,340
WTP using medians trimmed data (aggregated approach)	0.08	580,061	0.81	4,299	0.55	18,648

The formal tests of our hypotheses give the following results.

Hypothesis 1. We observe α_1 to be significantly lower than α_2 ($p<0.01$), which is consistent with our prediction resulting from loss aversion. Related to this finding, the estimated WTP is significantly higher for TTO1 than for TTO2 under both the additive and the multiplicative model ($p<0.01$), indicating individuals are willing to give up more lifetime to move from a subsistence income to current income, than to move from their current income to a luxury income.

Hypothesis 2. X_1 is significantly higher than X_3 ($p<0.01$), indicating that loss aversion is stronger for life duration than for QoL.

5. Discussion

This research set out to explore a novel method to value life years by means of trading life years for income. We applied three different procedures to elicit WTP with this method under different assumptions about the utility functions for health and money. Our trimmed WTP estimates give numbers that are comparable to estimates found in the literature (Mason et al., 2009; Pinto-Prades et al., 2009; Shiroiwa et al., 2013), although the high variation across procedures indicates a high susceptibility to the particular procedure employed. Likewise, the differences between models show the large influence of the particular assumptions about the utility functions for life duration and consumption on WTP estimates. Regarding the former, we find a difference in WTP between two procedures in the direction predicted by prospect theory. Furthermore, we observe significantly less non-trading when using QoL instead of life duration as response scale, although this does not necessarily translate into higher WTP for a healthy life year.

Regarding the additive model, the mean number of traded life years and the WTP estimates in TTO2 were comparable to those reported by Tilling et al. (2014) (their TTO1), who used the same method[2]. However, our WTP estimates were higher in TTO1 (their TTO2). One difference in the designs that the higher and lower income values were determined by the respondents themselves in our study, whereas these values were given by the experimenters in Tilling et al. (2014). Another difference was that we asked respondents to consider living their remaining lifetime in their current health state, while Tilling et al. (2014) instructed respondents to assume to spend the remaining lifetime in full health. Although we corrected for the respondents' own health by taking their VAS score into account, this may nevertheless have caused differences. Finally, Tilling et al. (2014) used a direct matching procedure, whereas we employed a combination of bisection and matching. However, these differences hold for both versions, so it is not evident why we only observe higher WTP values for the gain version.

One of the limitations of this study was the high number of non-traders. Non-willingness to trade may be a sincere preference or an expression of protest against the nature of the exercise, but part of it may be the result of the magnitude of our trade unit. The minimum amount to be traded was 0.1 years, so if people were only prepared to sacrifice, say, 2 days, 0 was closer to this amount than 0.1 years. These respondents would then appear to have an infinite WTP, whereas in reality their WTP is finite (albeit high).

[2] However, they only estimated the additive model, so our comparisons only concerns that model.

Although many respondents did not trade at all, only about 25% of these non-traders expressed indifference between 10 years with the lower income and 10 years with the higher income, which would be the implication of non-trading. The other 75% preferred 10 years with the higher income, but picked the highest possible answer in the slider (i.e., 10 years with the higher income). Hence, it seems these respondents had some other reason to refuse any trading than being indifferent between earning a lower or a higher income. Explanations may be that they attempted to 'improve their position' or because their indifference value was between 9.9 and 10, which could not be expressed in our questionnaire. Future research may therefore experiment with other designs such as only presenting binary choices or not applying sliders.

The difference in non-trading behaviour between TTO1 and TTO2 may also have been caused by the amount of difference between current and luxury income, versus the difference between subsistence and current income, which of course differed between subjects. The former difference was higher on average than the latter. Consequently, respondents were more likely to give up lifetime in the current-luxury trade-off than in the subsistence-current trade-off.

A second limitation was that a substantial part of the respondents trade too many life years, causing negative WTP for a life year. This finding may be caused by respondents not seriously engaging in or comprehending the task (despite our explanation of the fact that their answer implies their total income will be lower and their life span shorter), but also to a sincere preference for a high income per period. Obviously, a negative WTP is nonsensical, as it implies these people would not want to live an additional year in full health even if it would cost them no money at all. However, a possible rationalization for this behavior might be that individuals derive such a high amount of utility from having a high income *per month* that they prefer a short life with a high monthly income over a longer life with more *total income* but a lower *monthly income*. This argument would translate into a composite utility function that incorporates utility of income instead of utility of wealth. More research is required to sort out this question.

Third, our results reveal that respondents tend to pick the highest amount of the range in the scroll bar question, resulting in a multi-peaked answer distribution. This observation points toward some kind of preference construction, where respondents are influenced by the initial question. That is, they may be subject to an anchoring bias, as reported earlier in TTO and WTP studies (Samuelsen et al., 2012; Ternent and Tsuchiya, 2013; van Exel et al., 2006).

Furthermore, their indifference value may not necessarily represent a true indifference, but instead a wish of subjects to improve their position.

Fourth, the TTO2 and QTO versions generated significantly fewer respondents who were indifferent between 10 years with income L [C] and 10 years with income C [S], or who even preferred the latter to the former option, than the TTO1 version. Given that TTO2 and QTO were always asked after TTO1, this finding could be due to a learning effect. Future research randomizing the order of these tasks is needed to test this possibility.

Notwithstanding these shortcomings, several conclusions and areas for future research emerge from our experiment. First, WTP is sensitive to both the amount of the income compared and to the currency used to trade off health for money (i.e. duration or QoL). Second, large differences in WTP result from making different assumptions regarding the lifetime utility function, stressing the need to obtain a valid measurement of the parametric shape of this function. Third, the high numbers of infinite and negative WTP estimates indicate that the investigated procedure also has drawbacks (like common WTP approaches). The presence of non-traders is inherent to the WTP and TTO approaches in general and hard to resolve. The presence of over-traders is specific to the current method.

Our findings were consistent with Hypothesis 1 (i.e., $\alpha_1<\alpha_2$), but loss aversion need not be the only reason for this. One other possibility would be that the multiplicative model is valid but that it needs to be accompanied by a nonlinear utility function over life years (Abellán-Perpinán et al., 2006; Attema et al., 2012; Bleichrodt and Pinto, 2005; Wakker and Deneffe, 1996). For instance, if individuals discount the future, this reflects a concave utility of life duration function (e.g. a power function with power smaller than 1). The power estimates of the utility function over income may turn out to be constant across questions if we allow for such a generalisation, indicating our rejection is due to an invalid assumption regarding the utility of life duration. This emphasizes the importance of controlling for both utility functions. In addition, the multiplicative model may be valid with a linear utility of life duration, but with the utility function for income having another parametric shape than one belonging to the power family. Its shape may instead be exponential, reflecting constant absolute risk aversion instead of constant relative risk aversion. However, applying an exponential function is more elaborative as it does not give an analytical solution for the exponent and has to be solved numerically for each respondent. In sum, our findings do neither necessarily reject the multiplicative or additive shapes of the utility of health and wealth, nor do they necessarily imply the presence of loss aversion; they only indicate that it is inappropriate to model the responses by a combination of a linear utility of life duration function, a power function of wealth, and the assumption of no loss aversion. Further research is required to test which parametric shape best fits lifetime preferences and whether assuming prospect theory causes an improvement in the descriptive validity of individual behavior.

The significant difference between X_1, the answer to TTO1, and X_3, the answer to QTO, rejects Hypothesis 2 (i.e., $\lambda_H=\lambda_Q$), and implies a violation of the QALY model. The sign of the difference implies more loss aversion with respect to life duration than with respect to QoL. This finding is consistent with the tendency of people to refuse trading off life years in classical TTO (Arnesen and Trommald, 2005). However, WTP is only higher for TTO1 than QTO under the additive model; in fact, WTP is significantly lower for TTO1 than for QTO when assuming the multiplicative model. The major reason for these contradictory findings seems to be the large number of respondents with negative WTP: for QTO, negative WTP was only possible under the multiplicative model, but not under the additive model, resulting in much lower median WTP estimates under the multiplicative model for this procedure. This highlights the importance of the underlying lifetime utility function.

This research clearly has an explorative character. Nevertheless, given the existing methodological problems with traditional WTP, new approaches should be developed and explored. Furthermore, as described earlier in this discussion, our results open up several new and important areas for future research.

Appendix – Mathematical derivations

<u>Estimation of α in TTO1 in case of loss aversion:</u>

$$10 \times Q \times C^{\alpha} = [X_1 - \lambda_H(10 - X_1)] \times Q \times L^{\alpha}$$

$$\left(\frac{C}{L}\right)^{\alpha} = \frac{X_1 - (10 - X_1)\lambda_H}{10} \Leftrightarrow \left(\frac{C}{L}\right)^{\alpha} = \frac{X_1(1 + \lambda_H)}{10} + \lambda_H \Leftrightarrow \alpha = \frac{ln\left(\frac{X_1(1 + \lambda_H)}{10} + \lambda_H\right)}{\ln(C) - \ln(L)}$$

<u>Estimation of α in TTO2 in case of loss aversion:</u>

$$10 \times Q \times [S - \lambda_M(C - S)]^{\alpha} = [X_2 - \lambda_H(10 - X_2)] \times Q \times C^{\alpha}$$

$$\left(\frac{(1 + \lambda_M)S - C\lambda_M}{C}\right)^{\alpha} = \frac{X_2 - (10 - X_2)\lambda_H}{10} \Leftrightarrow \left(\frac{(1 + \lambda_M)S - C\lambda_M}{C}\right)^{\alpha} = \frac{(1 + \lambda_H)X_2 - 10\lambda_H}{10} \Leftrightarrow \left(\frac{(1 + \lambda_M)S}{C} - \lambda_M\right)^{\alpha} = \frac{(1 + \lambda_H)X_2}{10} - \lambda_H \Leftrightarrow \alpha = \frac{ln\left(\frac{(1 + \lambda_H)X_2}{10} - \lambda_H\right)}{ln\left(\frac{(1 + \lambda_M)S}{C} - \lambda_M\right)}$$

<u>Estimation of α in QTO in case of loss aversion:</u>

$$10 \times Q \times C^{\alpha} = 10 \times \left[\frac{X_3}{10} - \lambda_Q(1 - \frac{X_3}{10})\right] \times Q \times L^{\alpha}$$

$$\left(\frac{C}{L}\right)^{\alpha} = \frac{X_3 - (10 - X_3)\lambda_Q}{10} \Leftrightarrow \left(\frac{C}{L}\right)^{\alpha} = \frac{X_3\left(1 + \lambda_Q\right)}{10} + \lambda_Q \Leftrightarrow \alpha = \frac{ln\left(\frac{X_3\left(1 + \lambda_Q\right)}{10} + \lambda_Q\right)}{\ln(C) - \ln(L)}$$

References

-Abellán-Perpinán JM, Pinto JL, Méndez-Martinez I, Badia-Llach X. Towards a better QALY model. Health Economics 2006;15; 665-676.

-Abelson P. The Value of Life and Health for Public Policy. Economic Record 2003;79; S2-S13.

-Arnesen T, Trommald M. Are QALYs based on time trade-off comparable?--A systematic review of TTO methodologies. Health Economics 2005;14; 39-53.

-Attema AE, Brouwer WBF, l'Haridon O. Prospect theory in the health domain: A quantitative assessment. Journal of Health Economics 2013;32; 1057-1065.

-Attema AE, Bleichrodt H, Wakker PP. A direct method for measuring discounting and QALYs more easily and reliably. Medical Decision Making 2012;32; 583-593.

-Bleichrodt H, Pinto JL. The validity of QALYs under non-expected utility. The Economic Journal 2005;115; 533-550.

-Bleichrodt H, Pinto JL. Loss aversion and scale compatibility in two-attribute trade-offs. Journal of Mathematical Psychology 2002;46; 315-337.

-Bleichrodt H, Pinto JL, Abellán-Perpinán JM. A consistency test of the time trade-off. Journal of Health Economics 2003;22; 1037-1052.

-Bobinac A, van Exel NJA, Rutten FFH, Brouwer WBF. Willingness to Pay for a Quality-Adjusted Life-Year: The Individual Perspective. Value in Health 2010;13; 1046-1055.

-Bobinac A, van Exel J, Rutten FFH, Brouwer WBF. GET MORE, PAY MORE? An elaborate test of construct validity of willingness to pay per QALY estimates obtained through contingent valuation. Journal of Health Economics 2012;31; 158-168.

-Bobinac A, van Exel J, Rutten FFH, Brouwer WBF. The Value of a QALY: Individual Willingness to Pay for Health Gains Under Risk. PharmacoEconomics 2014;32; 75-86.

-Claxton K, Paulden M, Gravelle H, Brouwer W, Culyer AJ. Discounting and decision making in the economic evaluation of health- care technologies. Health Economics 2011;20; 2-15.

-Claxton, K., Walker, S., Palmer, S., Sculpher, M. 2010. Appropriate perspectives for health care decisions.

-College Voor Zorgverzekeringen (Dutch board of insurers). Dutch guidelines for pharmacoeconomic researchwww.ispor.org/PEguidelines/source/PE_guidelines_english_Netherlands.pdf. December 22013.

-Culyer A, McCabe C, Briggs A, Claxton K, Buxton M, Akehurst R, Sculpher M, Brazier J. Searching for a threshold, not setting one: the role of the National Institute for Health and Clinical Excellence. Journal of Health Services Research & Policy 2007;12; 56-58.

-Dalmau-Matarrodona E. Alternative approaches to obtain optimal bid values in contingent valuation studies and to model protest zeros. Estimating the determinants of individuals' willingness to pay for home care services in day case surgery. Health Economics 2001;10; 101-118.

-Dolan P, Edlin R. Is it really possible to build a bridge between cost-benefit analysis and cost-effectiveness analysis? Journal of Health Economics 2002;21; 827-843.

-Domeij D, Johannesson M. Consumption and Health. The B.E.Journal of Macroeconomics 2006;6; 1-30.

-Eeckhoudt L, Godfroid P, Marchand M. Risque de santé, médecine préventive et médecine curative. Revue D'Economie Politique 1998;108; 321-337.

-Finkelstein A, Luttmer EFP, Notowidigdo MJ. WHAT GOOD IS WEALTH WITHOUT HEALTH? THE EFFECT OF HEALTH ON THE MARGINAL UTILITY OF CONSUMPTION. Journal of the European Economic Association 2013;11; 221-258.

-Frew EJ, Whynes DK, Wolstenholme JL. Eliciting Willingness to Pay: Comparing Closed-Ended with Open-Ended and Payment Scale Formats. Medical Decision Making 2003;23; 150-159.

-Gyrd-Hansen D. Willingness to pay for a QALY. Health Economics 2003;12; 1049-1060.

-Gyrd-Hansen D, Kjær T. Disentangling WTP per QALY data: different analytical approaches, different answers. Health Economics 2012;21; 222-237.

-Hackl F, Pruckner GJ. Warm glow, free-riding and vehicle neutrality in a health-related contingent valuation study. Health Economics 2005;14; 293-306.

-Hayes K, Lockwood M, Anderson G. Estimating the Benefits of Water Quality Improvements in the Upper Narragansett Bay. Marine Resource Economics 1992;7; 75-85.

-Hirth RA, Chernew ME, Miller E, Fendrick AM, Weissert WG. Willingness to Pay for a Quality-adjusted Life Year: In Search of a Standard. Medical Decision Making 2000;20; 332-342.

-Johannesson M, Johansson P. Is the valuation of a QALY gained independent of age? Some empirical evidence. Journal of Health Economics 1997;16; 589-599.

-Johannesson M, Meltzer D. Editorial: Some reflections on cost-effectiveness analysis. Health Economics 1998;7; 1-7.

-Johnson FR, Desvousges WH, Ruby MC, Stieb D, De Civita P. Eliciting Stated Health Preferences: An Application to Willingness to Pay for Longevity. Medical Decision Making 1998;18; S57-S67.

-Kahneman D, Tversky A. Prospect Theory: An Analysis of Decision under Risk. Econometrica 1979;47; 263-291.

-King JT, Tsevat J, Lave JR, Roberts MS. Willingness to Pay for a Quality-Adjusted Life Year: Implications for Societal Health Care Resource Allocation. Medical Decision Making 2005;25; 667-677.

-Lundberg L, Johannesson M, Silverdahl M, Hermansson C, Lindberg M. Quality of life, health-state utilities and willingness to pay in patients with psoriasis and atopic eczema. The British Journal of Dermatology 1999;141; 1067-1075.

-Mason H, Jones-Lee M, Donaldson C. Modelling the monetary value of a QALY: a new approach based on UK data. Health Economics 2009;18; 933-950.

-Moffett ML, Suarez-Almazor M. Prospect theory in the valuation of health. Expert Rev Pharmacoeconomics Outcomes Res 2005;5; 499-505.

-NICE. Guide to the methods of technology appraisalhttp://www.nice.org.uk/aboutnice/howwework/devnicetech/technologyappraisalprocessguides/guidetothemethodsoftechnologyappraisal.jsp. 20 February2012.

-O'Brien B, Drummond M. Statistical Versus Quantitative Significance in the Socioeconomic Evaluation of Medicines. PharmacoEconomics 1994;5; 389-398.

-Olsen JA, Donaldson C, Pereira J. The insensitivity of `willingness-to-pay' to the size of the good: New evidence for health care. Journal of Economic Psychology 2004;25; 445-460.

-Pennington M, Baker R, Brouwer W, Mason H, Hansen DG, Robinson A, Donaldson C, the EuroVaQ Team. COMPARING WTP VALUES OF DIFFERENT TYPES OF QALY GAIN ELICITED FROM THE GENERAL PUBLIC. Health Economics 2013; n/a-n/a.

-Pinto-Prades JL, Loomes G, Brey R. Trying to estimate a monetary value for the QALY. Journal of Health Economics 2009;28; 553-562.

-Robinson A, Gyrd-Hansen D, Bacon P, Baker R, Pennington M, Donaldson C. Estimating a WTP-based value of a QALY: The 'chained' approach. Social Science & Medicine 2013;92; 92-104.

-Samuelsen CH, Augestad LA, Stavem K, Kristiansen IS, Rand-Hendriksen K. Anchoring bias in the Lead-Time Time Time Trade-Off. 29th EuroQoL plenary. 2012.

-Shiroiwa T, Igarashi A, Fukuda T, Ikeda S. WTP for a QALY and health states: More money for severer health states? Cost Effectiveness and Resource Allocation 2013;11; 22.

-Sloan FA, Viscusi WK, Chesson HW, Conover CJ, Whetten-Goldstein K. Alternative approaches to valuing intangible health losses: The evidence for multiple sclerosis. Journal of Health Economics 1998;17; 475-497.

-Smith RD. The relative sensitivity of willingness-to-pay and time-trade-off to changes in health status: an empirical investigation. Health Economics 2001;10; 487-497.

-Starmer C. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. Journal of Economic Literature 2000;38; 332-382.

-Stewart JM, Eamon O, Donaldson C, Shackley P. Do ordering effects matter in willingness-to-pay studies of health care? Journal of Health Economics 2002;21; 585-599.

-Tengstam S. DISABILITY AND MARGINAL UTILITY OF INCOME: EVIDENCE FROM HYPOTHETICAL CHOICES. Health Economics 2014;23; 268-282.

-Ternent L, Tsuchiya A. A Note on the Expected Biases in Conventional Iterative Health State Valuation Protocols. Medical Decision Making 2013;33; 544-546.

-Tilling C, Krol M, Attema AE, Tsuchiya A, Brazier J, van Exel NJA, Brouwer WBF. Exploring a new method for deriving the monetary value of a QALY. SSRN Working paper. 2014. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2539653.

-Treadwell JR, Lenert LA. Health Values and Prospect Theory. Medical Decision Making 1999;19; 344-352.

-Tversky A, Kahneman D. Advances in Prospect Theory: Cumulative Representation of Uncertainty. Journal of Risk and Uncertainty 1992;5; 297-323.

-van Exel NJA, Brouwer WBF, van den Berg B, Koopmanschap MA. With a little help from an anchor: Discussion and evidence of anchoring effects in contingent valuation. The Journal of Socio-Economics 2006;35; 836-853.

-Viscusi WK, Evans WN. Utility Functions That Depend on Health Status: Estimates and Economic Implications. The American Economic Review 1990;80; 353-374.

-Wakker P, Deneffe D. Eliciting von Neumann-Morgenstern Utilities When Probabilities Are Distorted or Unknown. Management Science 1996;42; 1131-1150.