Fair interpersonal utility comparison in the context of health valuation studies: early results of a new multi-step preference aggregation procedure

Paul Schneider*[,1], Ben van Hout[1], John Brazier[1]

[1] *ScHARR, University of Sheffield, UK*

Abstract

Background

The time trade-off (TTO) method is widely used to elicit individual preferences over (EQ-5D 3L, 5L, and Y) health states. Resulting utility values, measured on a scale that is anchored at full health and dead, are commonly aggregated across individuals, in order to derive a social value set. Here, we argue that this aggregation procedure is problematic, because it fails to take into account that TTO utility values are not only determined by an individual's preference for the quality, but also for the quantity of life. As a consequence, when health state utilities are aggregated across different individuals, individuals with a wider range of utility values have – simply for arithmetic reasons – more weight in the estimation of the social value set. Inspired by relative utilitarianism, we propose an alternative aggregation procedure, to equalise individuals' weights in the estimation of social health state values. We apply the alternative procedure to an EQ-5D 3L data set and compare the results against the conventional social tariff.

Methods

We used UK EQ-5D 3L data from the 1993 measurement and valuation of health study. As far as feasible, we replicated the main effects regression model and estimated social tariffs using the conventional and the proposed alternative utility aggregation procedure. Differences between both approaches were assessed by comparing model coefficients, as well as absolute and relative social health state values. In addition, we used linear and polynomial regression models to investigate to what extent individuals' utility ranges are associated with influence on the (relative) social tariff.

Results

2,985 participants were included in the analysis. The alternative aggregation procedure yielded different results than the conventional method: The main effects model coefficients differed between 0% and 8%, and social values were generally higher (mean error = 0.022) in the alternative tariff, and 156 (64.2%) health states had a different (relative) rank. For the conventional method, we found a strong, non-linear relationship between an individuals' utility ranges and their influence on the (relative) social value set (R2=0.18), while for the alternative method, the relationship was much weaker (R2=0.04).

Conclusion

The conventional TTO preference aggregation procedure gives individuals with a wider range of utility values systematically more weight in the estimation of the EQ-5D 3L social tariff. The proposed alternative method can resolve this issue and provides significantly different social health state values.

*School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, 30 Regent St, Sheffield S1 4DA, UK. Email: p.schneider@sheffield.ac.uk

1 Introduction

The measurement and valuation of health-related quality of life (HRQOL) is an integral component of health economic evaluation [1, 2]. The methods that are being used to value changes in the quality and the quantity of life due to health interventions have evolved significantly over the last decades. However, while much attention has been paid to the psychometric and statistical properties, little consideration was given to their normative underpinnings and implications. In fact, a coherent theoretical basis for the measurement and valuation of health is missing.

In previous preliminary work we have argued that the aggregation of health state utilities derived from time trade-off method (TTO) - which is the most widely used method to elicit health state preferences - might yield *unfair* social outcomes [3]. This study aims to extend our theoretical work by testing the proposed solution empirically using the UK social EQ-5D 3L value set.

The remainder of this paper is divided into three sections. In the next section (2), we outline our theoretical framework, explain our criticism of the conventional TTO method, and describe an alternative utility aggregation mechanism. In section 3, we report on an empirical application of the proposed alternative method on EQ-5D 3L valuation data from the UK. Finally, empirical and theoretical findings are discussed together in section 4.

Glossary of key concepts

Health state utilities: preference intensities, measured on a cardinal scale anchored at full health (=1) and dead (=0). May include negative values for states worse than dead.

Social tariff: a function which maps health states to social values, representing the collective preference of a group of individuals.

Relative social tariff: a function which maps health states to relative social values (the value of one state compared to all others), normalised between the best (=1) and the worst health state (=0).

Utility range: An individual's range of utility values, i.e. the utility distance between the best and the worst health state.

Influence: The effect that an individual's preference function has on the (relative) social tariff, given the preferences of all other individuals.

2 Theoretical framework

2.1 The conceptual basis of social value sets

A fundamental problem in health economics is that HRQOL cannot be measured (purely) objectively. It is, to some extent, a social construct. Measuring it requires making subjective value judgements [4]. The EQ-5D and other instruments that aim to measure HRQOL are therefore based on individual preferences. Those commonly elicited through TTO exercises.

The TTO method measures preferences by identifying points of indifference between shorter periods in full health and usually a fixed period in an ill-health state [5, 6]. The resulting health state utility values lie on a scale with full health and being dead as anchor points at one and zero respectively. States considered worse than being dead have negative values, which – purely for practical reasons - have a lower bound of minus one [7].

When different individuals are asked to state their preferences over a number of health states, there can be considerable disagreement between individuals (is mild pain better or worse than severe mobility impairment?). To determine the social value of a health state, which is supposed to represent the preference of the group of individuals as a whole, some form of compromise must be reached [4, 8]. To this end, utilities are aggregated across individuals, usually by taking the arithmetic average [9]. When cardinal aggregation procedures, such as the average, are applied, utility values of different individuals are implicitly assumed to be directly comparable with each other - averaging over values that are measured on different scales, such as degrees Celsius and degrees Fahrenheit, would simply not be meaningful [10].

On closer inspection, however, TTO health state utilities might not be interpersonally comparable. Even though they appear to be measured with a common yardstick, the units of measurement might differ between individuals.

2.2 Interpersonally (in)comparable health state utilities

In a previous preliminary work, we have discussed the conceptual problems with the TTO utility scale and their implications in more detail [3]. Here, we confine ourselves to present the main argument. It starts from the premise that there is no scientific way to quantify and compare the intensity of preferences between different individuals. The measurement of health state preferences on a 1-0 scale between full health and dead can thus not be assumed to capture utilities (or HRQOL) in some natural units. Rather, the TTO scale is imposed on normative grounds, as a matter of fairness. The value of one year in full health is set to be 1 for all individuals because everyone's preferences should, in principle, carry equal weight in the determination of the social tariff. Whether some individuals are able to enjoy full health more than others are not considered relevant.

The TTO utility aggregation procedure has some similarities with the concept of relative utilitarianism, which normalises every individual's utility scale between their best (=1) and their worst alternative (=0) [11]. This method seems intuitively attractive, as it expresses the democratic ideal that each individual is "to count for one and no more than one" [12, p. 283]. The approach has also been widely discussed in the fields of social choice theory and welfare economics - see for example Hausman [13, p. 489], who argues (although for different reasons) that relative utilitarianism is the only "correct way of making interpersonal comparisons".

In contrast to relative utilitarianism, however, the TTO method does not assign a zero value to the worst state, but to 'being dead' - which we consider to be qualitatively different from a health state. This has considerable implications for the comparability of health state utilities between individuals. As a consequence, the effective range of utility values, that is the distance between the best and the worst health state, can differ considerably between individuals. In fact, utility ranges can vary between zero (from 1 to 1), in individuals who refuse to trade any survival time, and two (from 1 to -1), in individuals who are very averse to poor health states. In general, the more dis-utility an individual derives from poor health and/or the more will-

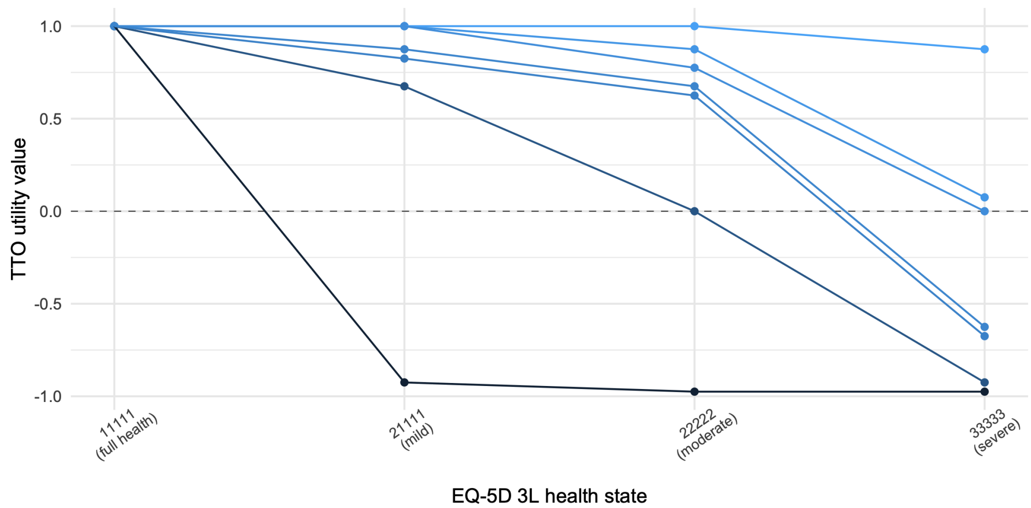

Figure 1: Differences between individuals' utility ranges. Shown are the health state utility values of a selected sample of eighth individuals. Each line represents the preferences of one individual over four EQ-5D 3L health states. Data: Williams *et al.* [16].

ing they are to trade survival time (in full health) for gains in HRQOL, the wider their utility range. This raises the question whether TTO utility values have the same meaning for different individuals.

In empirical health valuation studies it can also be seen that while some individuals are willing to give up a large proportion of their remaining life time for relatively minor gains in HRQOL, others only trade small proportions, and some even refuse to trade any survival time at all. Moreover, some individuals consider certain states worse than dead, while others do not [14, 15]. See figure 1 for an illustration of the differences between individuals' utility ranges.

2.3 Utility range and influence

Differences in utility ranges between individuals are potentially problematic, because individuals with a wider utility range have – simply for arithmetic reasons – more leverage in the estimation of the (average) social value of a health state. More specifically, if social values are being estimated for more than one health state, individuals with a wider utility range have more influence on the relative cardinal ordering of health states. This means, they have disproportionate say over the question how much better or worse one year in state i is compared to one year in state j.

These differences in influences between individuals should be considered illegitimate, because they are not due to genuine differences in the evaluation of the HRQOL, but are caused by differences in the preferences for survival time, relative to HRQOL - One may think of this as the exchange rate between units of quality and units of quantity of life. There does not seem to be any compelling reason, why individuals who are generally more willing to trade survival time (in full health) for gains in HRQOL should have more say in the social tariff.

To illustrate the point, consider a utility comparison between two individuals: Alice and Bob. Suppose that for Alice, all health states have TTO values between 1 and 0.9, while for Bob, they range between 1 and -1. Because of the wider utility range, Bob has more say in the relative social value set, when utilities are aggregated. We argue that this should be considered unfair, because from a narrow utility range, it does not follow that Alice only has weaker preferences over the health states. Instead, she might just have a higher rate of substitution between quality and quantity of life: her opportunity costs for giving up lifetime might be higher, for example because she has dependents to care for. It would then be illegitimate to give her less weight in the social tariff than Bob.

To avoid this type of bias, we propose an alternative aggregation procedure, which aims to equalise individuals' influences in the estimation of the relative social tariff.

2.4 Model

Inspired by relative utilitarianism [11], we propose a novel *Multi-step Utility Aggregation Procedure* (MUAP) for TTO values. The procedure consists of four steps, which are outlined below. Subsequently, we provide a simple example to illustrate how the method works. For convenience, we first introduce some basic notations.

Let $H \in \{h_1, h_2, \ldots, h_k\}$ denote a set of k health states, for which social values are to be determined for a group of n individuals. Individual j's health states preferences, denoted $u_j(H) \in \{u_j(h_1), u_j(h_2), \ldots, u_j(h_k)\}$, are measured on the TTO scale, anchored at full health (=1) and dead (=0). By definition, full health has a value of 1, which is everyone's highest utility value ($u(h_{full}) = 1 = \max u(H)$), while the lowest value can take different values with $1 > \min u_j(H) \geq -1$ (we have to define $\min u_j(H) < 1$ to avoid division by zero). Finally, j's utility range is given by $r_j = \max u(H) - \min u_j(H)$. Note that 'being dead' is not considered a health state.

The Conventional Aggregation Method (CAM)

The Conventional utility Aggregation Method (CAM) defines the social tariff function $F(.)$ simply as the average utility, as shown below.

$$F(H) = \frac{\sum_{j=1}^{n} u_j(H)}{n} \implies F(h_i) = \frac{\sum_{j=1}^{n} u_j(h_i)}{n}$$

Multi-step utility aggregation procedure (MUAP)

The MUAP is based on the idea that a social tariff consists of two main components, for which a social welfare function must be defined: a relative tariff, which determines the (relative) value of one health states compared to another, and a scaling factor, which anchors the relative tariff on the full-health-dead scale, and thereby determines the social rate of substitution between quality and quantity of life.

To derive a relative social tariff, everyone's utilities are normalised between their best (full health) and their worst health state. Normalised preferences and then aggregated by taking the average. The scaling factor is determined by the average utility range. The relative social tariff and the scaling factor are then combined to derive the alternative MUAP social tariff $S(.)$. A formal description of the four steps are given below.

The MUAP - Step-by-step instructions

1. **Min-Max Normalisation**
   For every individual $j \in n$, health state utilities are normalised between the highest value ($= 1 = u_j(h_{full})$, by definition) and the lowest value, so that everyone's utilities range from 1 (best) to 0 (worst).

$$u'_j(H) = \frac{u_j(H) - \min u_j(H)}{1 - \min u_j(H)} \quad (1)$$

2. **Relative social tariff**
   Normalised utility values are aggregated across individuals to derive the *relative* social tariff.

$$S'(H) = \frac{\sum_{j=1}^{n} u'_j(H)}{n} \quad (2)$$

3. **Scaling factor**
   Individual utility ranges ($r_j = 1 - \min u_j(H)$) are aggregated across individuals to derive the scaling factor.

$$R = \frac{\sum_{j=1}^{n} r_j}{n} \quad (3)$$

4. **Rescaling**
   Finally, the relative social tariff are re-scaled to the original full-health-dead scale, using the average scaling factor R, and re-anchored at full health with $1 - R$.

$$S(H) = R * S'(H) + 1 - R \quad (4)$$

Notations:
H: set of all health states $h_1, h_2, \ldots, h_k$
$u'_j(H)$: individual j's normalised preferences
$S'(.)$ relative social tariff
$u_j(H)$: individual j's health state preferences
$\min_j(H)$ / $\max_j(H)$: j's lowest/highest utility
$S(.)$ final alternative social tariff.

2.5 A simple example

To demonstrate the conceptual difference between the CAM and the MUAP, consider the example from above (see section 2.3). Suppose *Alice* and *Bob* state their preferences over four health states: h_{full} (full health); h_a (immobile); h_b (pain); h_0 (immobile and pain). Table 1 shows the utility values of both individuals and the corresponding CAM and MUAP social value sets. The results are visually illustrated in figure 2.

Alice and Bob both consider h_{full} the best and h_0 the worst state, yet they have divergent preferences for the other two states. Alice prefers h_a (0.967 vs 0.933), while Bob prefers h_b (-0.333 vs 0.333). When utilities are aggregated using the CAM, the social tariff indicates that, as a group, Alice and Bob prefer h_b (0.633) over h_a (0.317). However, Alice's relative preference for h_a (0.667 vs 0.333) is exactly as strong as Bob's relative preference for h_b (0.333 vs 0.667). It is only because Bob has a wider utility range than Alice (2.00 vs 0.1) that he has more influence on the social value set. When the MUAP is applied, influences are re-weighted and, as a group, Alice and Bob become indifferent between h_a and h_b (0.475 vs 0.475). The social values for h_{full} and h_0 are the same in both tariffs.

Table 1. Individual and social values

		h^*	h_a	h_b	h_0	r
CAM	u_{Alice}	1	0.967	0.933	0.900	
	u_{Bob}	1	-0.333	0.333	-1.000	
	F	**1**	**0.317**	**0.633**	**-0.050**	
MUAP	u'_{Alice}	1	0.667	0.333	0.00	0.100
	u'_{Bob}	1	0.333	0.667	0.00	2.00
	S' / $[R]$	1	0.500	0.500	0.000	[1.050]
	S	**1**	**0.475**	**0.475**	**-0.050**	

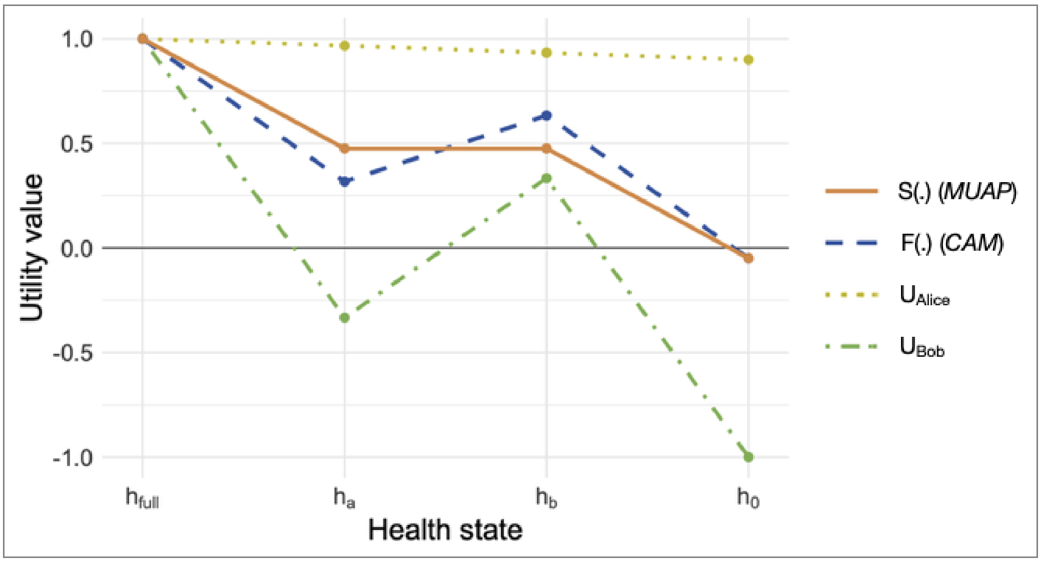

Figure 2: Shown are the two individual preference profiles (Alice = yellow; Bob = green) and the corresponding CAM (blue) and MUAP (orange) social value sets.

3 Empirical application

3.1 Objective

In order to test the proposed MUAP empirically, we applied it to an EQ-5D 3L health state valuation data set from the UK, and investigated the following two research questions:

1. What is the effect of using the MUAP, instead of the CAM, to derive a social EQ-5D 3L value set value set for the UK?
2. To what extent are individuals' utility ranges associated with individuals' influences on the CAM and MUAP relative social value sets?

3.2 Methods

3.2.1. The EQ-5D-3L Instrument

The EQ-5D 3L instrument is a generic preference-based instrument for the measurement and valuation of health. It consists of 243 health state, which are defined along five dimensions: mobility (MO), self-care (SC), usual activities (UA), pain/discomfort (PD), and anxiety/depression (AD). Each dimension has three levels: no problems (1), some problems (2), and severe problems (3). Health states can accordingly be referred to by a 5-digit code. '12321', for example, denotes a state with no mobility problems, some problems with self-care, unable to perform usual activities, some pain/discomfort, and no anxiety/depression; '11111' denotes full health; and '33333' denotes the (objectively) worst health state.

3.2.2. Setting and data

For this study, we used the UK EQ-5D 3L data set from the 1993 Measurement and Valuation of Health (MVH) Study [16–18]. We reproduced the methods of the original study as far as possible, but had to make two adjustments:

Firstly, the study sample originally consisted of 2,997 individuals, who were representative of the UK adult population. To avoid zero-division, we had to exclude 12 individuals who had a utility range of 0, i.e. who assigned a utility of 1 to all states.

Secondly, in the MVH, 42 of the 243 EQ-5D-3L health states were valued. Each respondent were asked to complete 12 TTO exercises, which included the worst health state '33333' and 11 other states. To be able to apply the MUAP, and to simplify the comparison with the CAM, we added 'full health' ('11111') with a utility of 1 as a 13th health state for each individual.

3.2.3. Statistical modelling of the CAM and MUAP social tariff

Since not all EQ-5D 3L health states were valued by every individual, a statistical model needs to be fitted to the data to estimate the average utility value for all 243 EQ-5D 3L health states. The CAM tariff is represented by the main effects model proposed by Dolan [17]. This well-known OLS regression model contains 12 coefficients, which are (relatively) easy to interpret: There are two dummy variables for each dimension (e.g. MO2 and MO3 for mobility), to represent the move between levels one (no problems) and two (some problems), and between two and three (severe problems). A constant (α) is subtracted for any move away from full health, and N3 represents an additional utility decrement for having severe problems on at least one dimension. The dependent variable was (one minus) the utility given to the health states. By convention, the social value for full health ('11111') is set to 1, although the statistical models would predict a lower value.

$$
\begin{aligned}
1-\hat{y} = \alpha + b_1 MO2 + b_2 MO3 + b_3 SC2 + b_4 SC3 + b_5 UA2 + b_6 UA3 + b_7 PD2 + \\
b_8 PD3 + b_9 AD2 + b_{10} AD3 + b_{11} N3 + \in
\end{aligned}
$$

To represent the alternative MUAP social tariff, we adjusted the main effects model corresponding to the steps described above (section 2.4): First, for each individual, utility values were Min-Max normalised between full health (=1) and the lowest utility value (=0), which could be assigned to '33333' or any other health state. Secondly, to estimate the relative social tariff, the main effects regression model was fitted with (one minus) the normalised/relative health state utility values as the dependent variable. Thirdly, the social scaling factor R was computed as the average individual utility range. Finally, to derive the alternative social value set, estimated relative social values for all health states were re-scaled, using the scaling factor.

The social scaling factor can also be applied to the MUAP model itself, in order to rescale the beta coefficient values. This allows estimating final social health state values directly, and it also enables a comparison of the relative importance of the explanatory variables between the CAM and the MUAP model.

3.2.4. Analysis

Differences between the CAM and the MUAP social value set were investigated in two ways. Firstly, we compared the coefficients of the CAM model with the rescaled coefficients of the MUAP model. Secondly, we compared the absolute social values for all 243 EQ-5D 3L health states, as well as their relative rankings.

Moreover, our theoretical framework predicts that, in the CAM, individuals with a wider utility range have more influence on the relative social value set, while in the MUAP, this association is not expected. In the absence of an established measure of 'influence', we defined it as the sum of changes in the social value set that occur due to the inclusion/exclusion of a particular individual. To compute an individual's influence on the relative social value sets, we re-fitted the CAM and MUAP main effects models without the individual in question, and assessed to what extent the estimates for the 243 relative social health state values change as a result. The sum of the absolute differences between the estimates with and without the individual in the model were then used as the measure of influence. The relationship between utility ranges and influences was assessed by visual inspection and and tested statistically using linear and polynomial regression models.

3.2.5. Data and code availability

The 1993 MVH data set is available on the UK Data Service [16]. The principal investigators gave us permission to publish the relevant data alongside the R code that was used to generate the results of this study on a data repository, where it is openly available for reuse and adaptation [19].

3.3 Results

3.3.1. Comparison between CAM and MUAP social tariffs

A total of 2,985 participants were included in the analysis. Table 2 shows the estimates for main effects models using the CAM and the MUAP approach. For most parameters, the differences were small. However, in the rescaled MUAP model, the beta for AD3 was 6% higher compared to the CAM model, and the beta for N3 was 8% lower, respectively.

Table 2. Main effects model - CAM and MUAP parameter estimates

	CAM	MUAP		$\frac{MUAP}{CAM}$
Variable	b (95%CI)	Normalised b (95%CI)	Rescaled b (95%CI)	Ratio
Intercept	0.05 (0.04; 0.06)	0.03 (0.03; 0.04)	0.05 (0.04-0.06)	1.01
MO2	0.07 (0.06; 0.08)	0.04 (0.04; 0.05)	0.07 (0.06-0.08)	0.99
MO3	0.24 (0.23; 0.25)	0.15 (0.14; 0.16)	0.24 (0.23-0.25)	1.01
SC2	0.12 (0.11; 0.13)	0.07 (0.06; 0.08)	0.11 (0.10-0.12)	0.96
SC3	0.11 (0.09; 0.12)	0.07 (0.06; 0.08)	0.11 (0.10-0.12)	1.05
UA2	0.04 (0.02; 0.05)	0.02 (0.02; 0.03)	0.04 (0.03-0.05)	1.04
UA3	0.06 (0.04; 0.07)	0.03 (0.03; 0.04)	0.06 (0.04-0.07)	1.01
PD2	0.13 (0.12; 0.14)	0.08 (0.07; 0.08)	0.12 (0.11-0.13)	0.96
PD3	0.26 (0.24; 0.27)	0.16 (0.15; 0.17)	0.26 (0.25-0.27)	1.00
AD2	0.08 (0.07; 0.10)	0.05 (0.04; 0.06)	0.08 (0.07-0.09)	0.97
AD3	0.16 (0.15; 0.18)	0.11 (0.10; 0.11)	0.17 (0.16-0.18)	1.06
N3	0.28 (0.27; 0.30)	0.16 (0.15; 0.17)	0.26 (0.25-0.28)	0.92
R^2	0.51	0.59	0.51	
Observations	38,805	38,805	38,805	

We estimated social EQ-5D 3L value sets for the UK using the CAM and the MUAP approach. The absolute social values and ranks for all 243 EQ-5D 3L health states are provided in table S1 in the appendix. A visual overview and comparison of the two social tariffs as a wholes are provided in Figure 3.

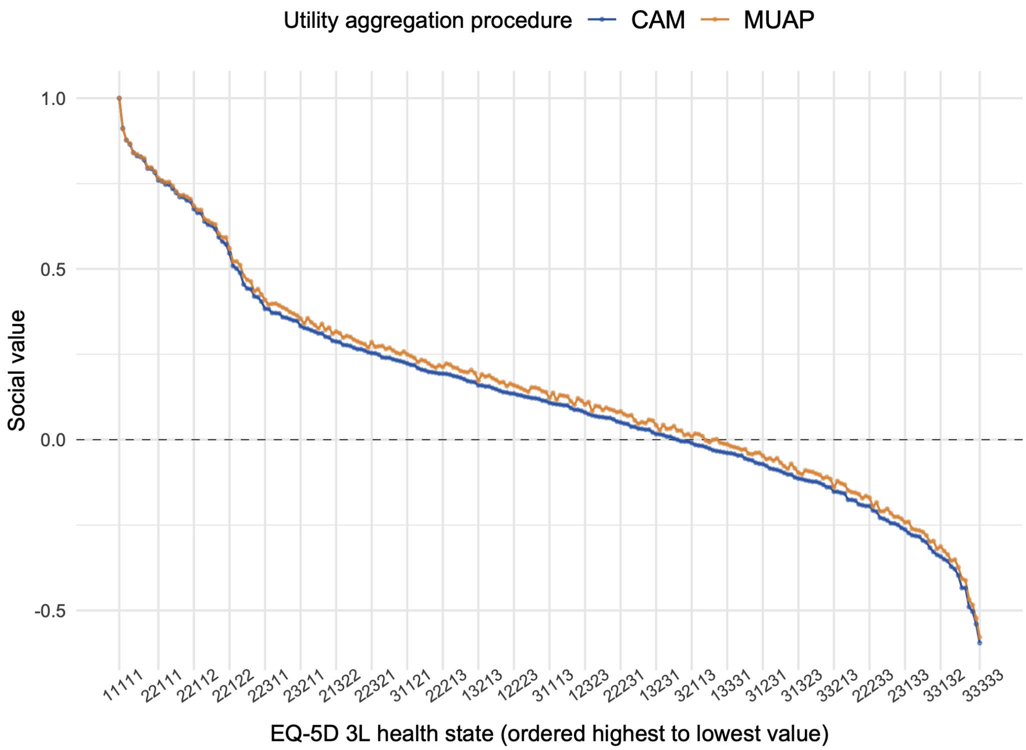

Figure 3: Comparison between the conventional CAM (blue line) and the alternative MUAP (orange line) UK social values for all 243 EQ-5D 3L health states - labels are shown for 25 states. Health states are arranged by their CAM social value on the x-axis in descending order.

The MUAP yielded different, and generally higher social values for EQ-5D 3L health states than the CAM: The mean (and mean absolute) difference between social values was 0.022 (SD = 0.008). Differences tended to be smaller for high and low ranked states, and larger for intermediate states. The maximum absolute difference of 0.036 was observed for state '22132' (CAM=0.003; MUAP=0.040). To provide a comparison, in the CAM, the average distance from one EQ-5D 3L health state to the next nearest state (rank i to $i+1$) was 0.007 (SD = 0.010).

The differences in absolute social values often resulted in changes of the ordering of health states. A total of 156 (64.2%) health states had a different (relative) rank in the MUAP tariff. The highest rank change of +6 (134→140) occurred for health state '31213'.

3.3.2. Utility ranges and influence

Individuals' utility ranges varied between 0.01 and 1.98, with an average (SD) of 1.62 (0.36) and a median (IQR) of 1.73 (0.55). Only 134 (4%) of the participants did not consider any health state to be worse than dead and had a utility range of less than 1. Furthermore, only 2,205 (74%) participants assigned the lowest utility value to the objectively worst health state '33333', while for 780 (26%) participants, at least one health state had a lower value.

Table 3 shows descriptive statistics for the measure of individuals' influences on the relative social tariff using the CAM and the MUAP approach. On average, the MUAP gave a single individual had less influence on the relative social value set than the CAM (0.024 vs 0.021). Measures of dispersion indicated that in the MUAP influences were also more evenly distributed. However, influence in the CAM was highly correlated with influence in the MUAP: The Spearman rank and Pearson correlation coefficients were 0.85 and 0.84, respectively.

Table 3. Dispersion of individuals' influences

	CAM	MUAP	$\frac{MUAP}{CAM}$
Mean	0.024	0.021	
Standard dev.	0.011	0.009	0.835
Minimum	0.004	0.004	
Maximum	0.092	0.064	
Range (max-min)	0.088	0.060	0.684
Max/min ratio	22.873	17.892	0.782
Q_{25}	0.016	0.015	
Median Q_{50}	0.022	0.019	
Q_{75}	0.029	0.026	
IQR (Q_{75}-Q_{25})	0.013	0.011	0.835

For both, the CAM and the MUAP, the association between between individuals' utility ranges and their influences on the relative social value set was modeled using linear, quadratic, and cubic regression models. Coefficient estimates are provided in table 4, and a visual illustration of the relationship is given in figure 3.

For the CAM, we found a clear, non-linear relationship between utility range an influence on the relative social values of health states. The three-degree polynomial model explained 18% of the variance of the influence between individuals. For the MUAP, model coefficients were also statistically significant (p¡.01) and showed similar patterns as for the CAM model. However, the effect was considerably smaller and, overall, individuals' utility ranges explained only a minor proportion of the variance (4%). In both approaches, the small group individuals with a narrow utility range of less than one (n=134) had high measures of influence.

Table 4. Association between utility ranges and influence on relative social tariffs - beta (95% CI)

	CAM linear	CAM square	CAM cubic	MUAP linear	MUAP sqaure	MUAP cubic
Intercept	0.02 (0.02; 0.02)	0.02 (0.02; 0.02)	0.02 (0.02; 0.02)	0.02 (0.02; 0.02)	0.02 (0.02; 0.02)	0.02 (0.02; 0.02)
Utility range	0.15 (0.13; 0.17)	0.15 (0.13; 0.17)	0.15 (0.13; 0.17)	-0.02 (-0.04; -0.01)	-0.02 (-0.04; -0.01)	-0.02 (-0.04; -0.01)
Utility range^2		0.18 (0.16; 0.20)	0.18 (0.16; 0.20)		0.09 (0.07; 0.11)	0.09 (0.07; 0.11)
Utility range^3			0.09 (0.07; 0.11)			0.04 (0.02; 0.06)
R^2	0.06	0.16	0.18	0.00	0.04	0.04
N	2,985	2,985	2,985	2,985	2,985	2,985

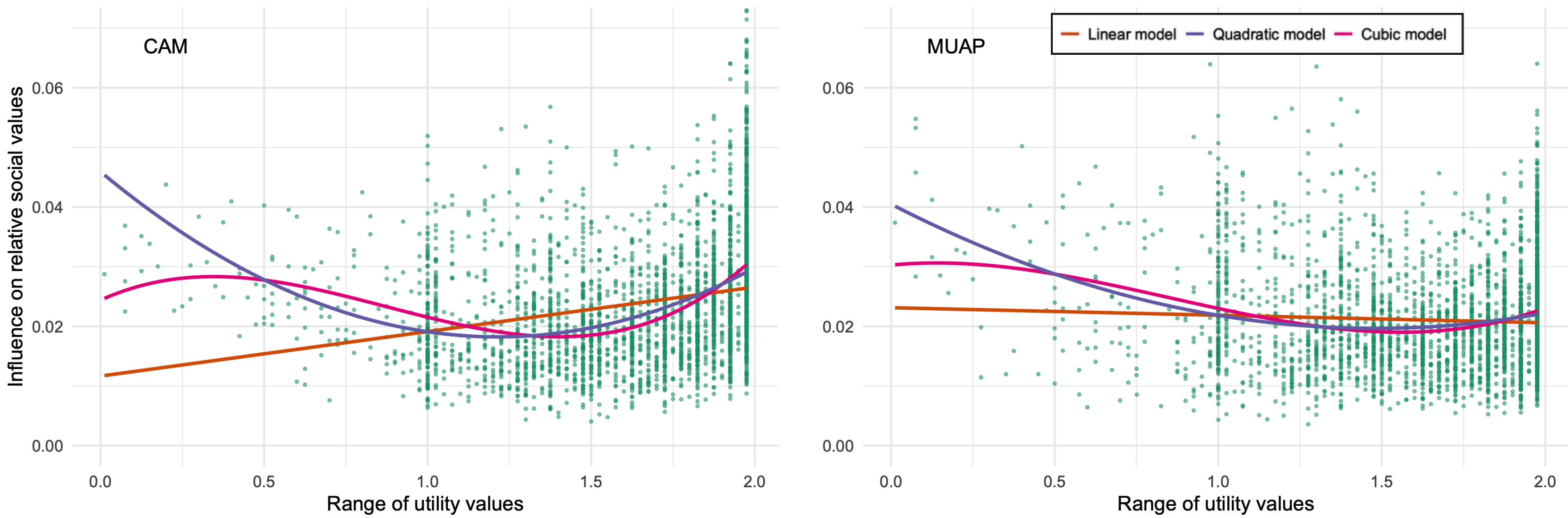

Figure 4: Relationship between individuals' utility ranges and their influences on the relative social tariff for the CAM (left) and the MUAP (right). Each point represents one participant (n=2,985). Lines show the linear (orange), quadratic (purple), and cubic model (magenta) estimates.

4 Discussion

4.1 Main findings

In this paper, we have outlined potential flaws in the procedure that is commonly used to aggregate individual TTO health state utilities into social value sets for use in economic evaluations. Our main theoretical argument is that the conventional method fails to ensure that every individual has an equal say in the social outcome. Instead, social value sets might predominantly represent the preferences of those individuals who are more willing to trade units of survival time for gains in units of HRQOL.

Using the well-known EQ-5D 3L data set of the MVH study [16–18], we found evidence that confirms our hypothesis. Individuals with a wider range of utility values have indeed systematically more weight in the estimation of the social EQ-5D 3L value set. Hoewver, the relationship was more complex than expected: not only did individuals with very high utility ranges have more influence, but also individuals with very low utility ranges. The group of individuals with low utility ranges was small (n = 134; 4%) though, and the effect might be due to the unusual preference profiles of these particular individuals.

Inspired by the concept of relative utilitarianism, we proposed an alternative utility aggregation method, the MUAP, to resolve the problem of the inequitable distribution of influence. The method equalises individuals' weights in the estimation of the relative social tariff, by normalising everyone's health state preferences between their best and worst state. We demonstrated the feasibility of the MUAP and were able to show that the it resolves the excess influence of individuals with wide utility ranges. Individuals with very low utility ranges still had more influence, but, as mentioned above, this is likely caused by their unusual preference profiles, and not a result of the low utility range itself.

The MUAP provided a different social EQ-5D 3L value set for the UK than the CAM. Health states values were generally higher in the MUAP and many states had a different relative rank (156; 64.2%). Even though differences between the MUAP and CAM social values might seem small for any given health state - even the maximum absolute difference was only 0.036 - in relation to 1-0 TTO scale, these differences should be considered relevant. If a social tariff, based on the MUAP, were used to value the health outcomes in economic evaluations, aggregate differences between QALY estimates could be significant. Even though this paper only focuses on EQ-5D and TTO, our approach should be equally applicable to other descriptive systems and the standard gamble method.

It should be noted that difference in the social tariff between the CAM and the MUAP can only be expected if individuals' utility ranges are associated with certain preferences for dimension weights. For example, individuals with a wider range might put more weight on mobility and less on pain/discomfort. The changes we observed in the beta coefficients between the CAM and the (rescaled) MUAP model provide some evidence for the existence of such associations: when the influence of individuals with wide utility ranges was equalised, the weight assigned to AD3 (severe anxiety/depression) increased by 6%, while the weight for N3 decreased by 8%. Unfortunately, we were unable to further investigate the underlying mechanism, since individual-level preference functions cannot be estimated reliably from the MVH data set.

4.2 Further considerations

The MUAP approach links directly to two other recently proposed (but unfortunately not yet fully published) alternative utility rescaling methods: Jakubczyk & Golicki [20] suggest to equalise utilities based on 'worst fears'. Their approach seems similar to ours, in that they normalise individuals' utilities between full health and either 'being dead' or the worst health state, depending on which has a lower utility. A whole series of alternative approaches to measure HRQOL were suggested by Sampson *et al.* [21]. They reject the use of 'being dead' in the TTO altogether and recommend a range of other outcomes, such as being unconscious, the worst health state, or minimum endurable HRQOL, which could instead be used in TTO exercises.

The principal difference between Jakubczyk & Golicki [20], Sampson *et al.* [21], the CAM, and our MUAP approach is the conception of 'being dead'. In the CAM, 'being dead' is defined as a health state and further assumed to be an absolute zero point for every individual; Jakubczyk & Golicki [20] also consider dead to be a health state, but they allow individuals to place a different value on it than zero; and Sampson *et al.* [21] argue that dead should not be considered relevant for the derivation of a social value set at all. Our MUAP takes yet another position. We do not consider 'being dead' to be a health state, and it is thus disregarded in the estimation of relative value of one health state compared to another. However, when determining the social rate of substitution between quality and quantity of life, 'being dead' is re-incorporated into the model as an anchor point at zero. Up until now, theoretical work on the significance of 'being dead' for the valuation of health has been scarce, and further research is required to better understand which of the conceptions of 'being dead' is most appropriate for use in health economic evaluations [3, 22, 23].

Our theoretical framework might also be helpful for addressing conceptual problems in the development of the EQ-5D-Y and other instruments for the valuation of health states in children [24]. Studies have shown that TTO values are higher for health states experienced by a 10-year-old child compared to health states experienced by an adult. At the same time, values derived through the VAS valuation techniques tend to be lower for states experienced by children [25]. These results are difficult to interpret and might seem paradoxical. Within our framework, however, they can be easily integrated. TTO values can be higher in children, even though their HRQOL is judged to be lower. This can be the case if the rate of substitution between quality and quantity of life is also higher. This means, people might be willing to trade fewer units of survival time, to gain a unit of HRQOL in in children than they are in adults. To harmonise the valuation systems for children and adults, it might be necessary to either apply a common social scaling factor, or increase the value of survival time in children, so that it can be worth more than 1 QALY.

4.3 Limitations

This study has several limitations which should be taken into account when interpreting our findings.

To be able to use the MUAP, we had to adjust the method of the original MVH study [18]. For the min-max-normalisation to be applicable, it is essential that the TTO values for the best and the worst health state are known for each individual. We assumed that in the context of the EQ-5D 3L descriptive system, it would not be necessary to elicit the preferences for all health states: full health has, by definition, the highest value (=1). The inclusion of full health as an additional health state for each individual is, however, problematic, because it complicates the interpretation of the constant in the main effects models, which is usually taken to represent any move away from full health [17]. For the worst state, we assumed that it was captured in the data set, since all participants valued health state '33333', which is objectively dominated by all other states. However, 780 (26%) participants assigned the lowest TTO value to one of the other 11 health states they valued (which we then used as the worst health state for the normalisation procedure). If we take this result at face value, and accept that individuals may consider another state to be worse than '33333', we would, in principle, have to elicit individuals' preferences for all states to ensure that the lowest TTO value is observed for everyone.

Finally, it is important to emphasise that our theoretical framework is only concerned with equalising the influences on the *relative* social tariff. To transform relative into absolute values, that is to rescale normalised utilities back to the full health-dead scale, we used a scaling factor. Even though this factor is also undoubtedly a social value and reflects some kind of societal compromise, it is not entirely clear what properties it has or whether comparisons of utility ranges between individuals are at all permissible.

4.4 Conclusion and outlook

This is the first application of an alternative, relative utilitarian method to aggregated TTO health state utilities across individuals into a social value set. We have outlined the theoretical advantages that the method has over the conventional approach and demonstrated its feasibility using an EQ-5D 3L dataset from the UK. Even though the differences in the resulting social health state values appear modest, in practice, the impact on aggregate QALY and incremental cost-effectiveness ratio estimates could be significant.

Even though several open questions remain, we think our theoretical framework provides a promising approach to generate more equitable social value estimates for use in health economic evaluations. More research is needed to better understand the practical and normative implications of this and other utility aggregation methods. Those methods should be tested empirically in other data sets and in other descriptive systems. Ideally, studies should seek to investigate the relationship between individual-level preference functions and the social value set.

To be able to decide which properties an aggregation method should or should not have, it seems crucial to establish a more sound theoretical foundation for the valuation of health in general. In this context, it is worth mentioning that in Welfare Economics, Social Choice Theory and parts of philosophy, the problem of interpersonal utility comparisons appears to have been addressed with much sophistication. In these fields, considerable effort has been devoted to rigorously investigate which types of utility comparisons are permissible, within and between individuals, for different sets of assumptions [10, 26]. We are convinced that research on the valuation of health would benefit from a closer consideration of this extensive body of work.

Acknowledgements

We are sincerely grateful to Koonal Shah, Robert Sugden, Simon McNamara, Nicolas Silva-Illanes and the participants of the 2020 HESG meeting in Newcastle for helpful discussions of the ideas in this paper. This work was supported by the Wellcome Trust Doctoral Training Centre in Public Health Economics and Decision Science (108903/Z/19/Z) and the University of Sheffield through a PhD scholarship. The usual disclaimer applies.

References

1. Karimi, M. & Brazier, J. Health, health-related quality of life, and quality of life: what is the difference? *Pharmacoeconomics* **34,** 645–649 (2016).
2. Brazier, J. Valuing health states for use in cost-effectiveness analysis. *Pharmacoeconomics* **26,** 769–779 (2008).
3. Schneider, P. & van Hout, B. Interpersonal comparability of health state utilities: why it is unfair to measure preferences in units of full-health-time, and what we can do about it - Working Paper. Available at: `https://bitowaqr.github.io/files/c2.pdf` (2020).
4. Devlin, N., Shah, K., Buckingham, K., *et al.* What is the normative basis for selecting the measure of 'average'preferences for use in social choices. *OHE research paper. Office of Health Economics, London* (2017).
5. Torrance, G. W. Social preferences for health states: an empirical evaluation of three measurement techniques. *Socio-economic planning sciences* **10,** 129–136 (1976).
6. Dolan, P., Gudex, C., Kind, P. & Williams, A. The time trade-off method: results from a general population study. *Health economics* **5,** 141–154 (1996).
7. Lamers, L. M. The transformation of utilities for health states worse than death: consequences for the estimation of EQ-5D value sets. *Medical care* **45,** 238–244 (2007).
8. Schneider, P. Social tariffs and democratic choice–do population-based health state values reflect the will of the people? (2019).
9. Xie, F., Gaebel, K., Perampaladas, K., Doble, B. & Pullenayegum, E. Comparing EQ-5D valuation studies: a systematic review and methodological reporting checklist. *Medical Decision Making* **34,** 8–20 (2014).
10. Fleurbaey, M. & Hammond, P. J. in *Handbook of utility theory* 1179–1285 (Springer, 2004).
11. Isbell, J. *Absolute games, in" Contributions to the Theory of Games, IV"(AW Tucker and RD Luce, Eds.)* 1959.
12. Rawls, J. *A theory of justice* (Harvard university press, 2009).
13. Hausman, D. M. The impossibility of interpersonal utility comparisons. *Mind* **104,** 473–490 (1995).
14. Jakubczyk, M., Golicki, D. & Niewada, M. The impact of a belief in life after death on health-state preferences: True difference or artifact? *Quality of Life Research* **25,** 2997–3008 (2016).
15. Augestad, L. A., Rand-Hendriksen, K., Stavem, K. & Kristiansen, I. S. Time trade-off and attitudes toward euthanasia: implications of using

'death'as an anchor in health state valuation. *Quality of Life Research* **22,** 705–714 (2013).

16. Williams, A. H., Gudex, C., Kind, P. & Dolan, P. *The measurement and valuation of health study* tech. rep. UK Data Service. SN: 3444. Available at: `http://doi.org/10.5255/UKDA-SN-3444-1` (1993).
17. Dolan, P. Modeling valuations for EuroQol health states. *Medical care,* 1095–1108 (1997).
18. MVH Group. The measurement and valuation of health: Final report on the modelling of valuation tariffs. *Centre for Health Economics, University of York.* Available at: `https://www.york.ac.uk/media/che/documents/reports/MVHFinalReport.pdf` (1995).
19. Schneider, P. P., van Hout, B. & Brazier, J. *Fair interpersonal utility comparison in the context of health valuation studies: data and code repository.* tech. rep. Available at: `http://github.com/bitowaqr/eq5d_muap` (2020).
20. Jakubczyk, M. K. & Golicki, D. What if 0 is Not Equal to 0? Interpersonal Utility Anchoring Using the Worst Fears. *Patient 12: 429. https://doi.org/10.1007/s40271-019-00368-9* (2019).
21. Sampson, C. J., Devlin, N. & Parkin, D. W. Drop dead: is anchoring at 'dead' a theoretical requirement in health state valuation? *Conference: 35th EuroQol Group Scientific Plenary* (2018).
22. Miyamoto, J. M., Wakker, P. P., Bleichrodt, H. & Peters, H. J. The zero-condition: a simplifying assumption in QALY measurement and multiattribute utility. *Management Science* **44,** 839–849 (1998).
23. Roudijk, B., Donders, A. R. T. & Stalmeier, P. F. Setting Dead at Zero: Applying scale properties to the QALY model. *Medical Decision Making* **38,** 627–634 (2018).
24. Kreimeier, S. & Greiner, W. EQ-5D-Y as a health-related quality of life instrument for children and adolescents: the Instrument's characteristics, development, current use, and challenges of developing its value set. *Value in Health* **22,** 31–37 (2019).
25. Rowen, D., Rivero-Arias, O., Devlin, N. & Ratcliffe, J. Review of Valuation Methods of Preference-Based Measures of Health for Economic Evaluation in Child and Adolescent Populations: Where are We Now and Where are We Going? *PharmacoEconomics,* 1–16 (2020).
26. d'Aspremont, C. & Gevers, L. Social welfare functionals and interpersonal comparability. *Handbook of social choice and welfare* **1,** 459–541 (2002).

Appendix

Table S1: Full UK social EQ-5D 3L value sets for the CAM and the MUAP.

Rank	Health state	CAM Value	MUAP Value	Difference Abs. (rank)
1	11111	1.000	1.000	0.00 (-)
2	11211	0.912	0.910	-0.00 (-)
3	21111	0.877	0.878	0.00 (-)
4	11112	0.865	0.867	0.00 (-)
5	21211	0.841	0.840	-0.00 (-)
6	12111	0.831	0.836	0.00 (-)
7	11212	0.828	0.829	0.00 (-)
8	11121	0.819	0.824	0.01 (-)
9	12211	0.794	0.797	0.00 (-)
10	21112	0.793	0.797	0.00 (-)
11	11221	0.782	0.786	0.00 (-)
12	22111	0.759	0.765	0.01 (-)
13	21212	0.757	0.759	0.00 (-)
14	21121	0.747	0.754	0.01 (+1)
15	12112	0.747	0.755	0.01 (-1)
16	11122	0.735	0.743	0.01 (-)
17	22211	0.723	0.727	0.00 (-)
18	21221	0.711	0.716	0.00 (+1)
19	12212	0.710	0.716	0.01 (-1)
20	12121	0.701	0.712	0.01 (-)
21	11222	0.698	0.705	0.01 (-)
22	22112	0.676	0.684	0.01 (-)
23	12221	0.664	0.673	0.01 (-)
24	21122	0.664	0.673	0.01 (-)
25	22212	0.639	0.646	0.01 (-)
26	22121	0.630	0.641	0.01 (-)
27	21222	0.627	0.635	0.01 (-)
28	12122	0.617	0.631	0.01 (-)
29	22221	0.593	0.603	0.01 (-)
30	12222	0.580	0.592	0.01 (+1)
31	11311	0.572	0.592	0.02 (-1)
32	22122	0.546	0.560	0.01 (-)
33	22222	0.509	0.522	0.01 (+1)
34	21311	0.501	0.522	0.02 (-1)
35	11312	0.489	0.511	0.02 (-)
36	12311	0.455	0.480	0.02 (-)
37	11321	0.443	0.468	0.03 (-)
38	13111	0.441	0.463	0.02 (-)
39	11113	0.419	0.434	0.01 (+1)
40	21312	0.417	0.441	0.02 (-1)
41	13211	0.404	0.425	0.02 (-)
42	22311	0.383	0.409	0.03 (-)
43	11213	0.383	0.396	0.01 (+2)
44	21321	0.371	0.398	0.03 (-)
45	12312	0.371	0.399	0.03 (-2)
46	23111	0.370	0.393	0.02 (-)
47	11322	0.359	0.387	0.03 (-)
48	13112	0.357	0.382	0.03 (-)
49	31111	0.353	0.374	0.02 (-)
50	13311	0.349	0.369	0.02 (-)
51	21113	0.348	0.364	0.02 (-)
52	23211	0.333	0.355	0.02 (+1)
53	11313	0.327	0.340	0.01 (+2)
54	12321	0.325	0.356	0.03 (-2)
55	13212	0.320	0.344	0.02 (-1)
56	31211	0.316	0.336	0.02 (+1)
57	21213	0.311	0.325	0.01 (+2)
58	13121	0.311	0.339	0.03 (-2)
59	12113	0.302	0.321	0.02 (+1)
60	22312	0.299	0.328	0.03 (-2)
61	11123	0.290	0.310	0.02 (+2)
62	21322	0.287	0.317	0.03 (-1)
63	23112	0.286	0.312	0.03 (-1)
64	23311	0.278	0.299	0.02 (+2)
65	11131	0.277	0.304	0.03 (-1)
66	13221	0.274	0.301	0.03 (-1)
67	31112	0.269	0.293	0.02 (-)
68	13312	0.265	0.288	0.02 (-)
69	12213	0.265	0.283	0.02 (+1)
70	31311	0.261	0.280	0.02 (+1)
71	21313	0.256	0.270	0.01 (+4)
72	22321	0.253	0.285	0.03 (-3)
73	11223	0.253	0.272	0.02 (+1)
74	23212	0.249	0.274	0.02 (-1)
75	12322	0.241	0.274	0.03 (-3)
76	11231	0.240	0.266	0.03 (+1)
77	23121	0.240	0.269	0.03 (-1)
78	32111	0.235	0.261	0.03 (-)
79	31212	0.233	0.255	0.02 (+1)
80	22113	0.230	0.251	0.02 (+1)
81	13122	0.227	0.258	0.03 (-2)
82	31121	0.223	0.250	0.03 (-)
83	13321	0.219	0.245	0.03 (-)
84	21123	0.218	0.240	0.02 (-)
85	12313	0.210	0.227	0.02 (+2)
86	21131	0.205	0.234	0.03 (-1)
87	23221	0.203	0.231	0.03 (-1)
88	32211	0.199	0.223	0.02 (-)
89	11323	0.198	0.216	0.02 (+3)
90	13113	0.196	0.211	0.01 (+5)
91	23312	0.194	0.218	0.02 (-)
92	22213	0.194	0.213	0.02 (+1)
93	11132	0.193	0.223	0.03 (-4)
94	13222	0.191	0.220	0.03 (-4)
95	31221	0.187	0.212	0.03 (-1)
96	11331	0.185	0.210	0.03 (-)
97	21223	0.182	0.201	0.02 (+1)
98	31312	0.177	0.199	0.02 (+1)
99	12123	0.172	0.197	0.03 (+1)
100	22322	0.170	0.204	0.03 (-3)
101	21231	0.168	0.195	0.03 (-)
102	13213	0.159	0.173	0.01 (+5)
103	12131	0.159	0.191	0.03 (-1)
104	11232	0.156	0.184	0.03 (-)
105	23122	0.156	0.188	0.03 (-2)
106	32112	0.151	0.180	0.03 (-1)
107	23321	0.148	0.175	0.03 (-1)
108	32311	0.143	0.167	0.02 (+1)
109	31122	0.139	0.169	0.03 (-1)
110	22313	0.138	0.157	0.02 (+2)
111	13322	0.135	0.164	0.03 (-1)
112	12223	0.135	0.159	0.02 (-1)
113	31321	0.131	0.156	0.02 (-)
114	33111	0.130	0.151	0.02 (+2)
115	21323	0.126	0.145	0.02 (+3)
116	23113	0.125	0.141	0.02 (+4)
117	12231	0.122	0.153	0.03 (-3)
118	21132	0.121	0.152	0.03 (-3)
119	23222	0.119	0.149	0.03 (-2)
120	32212	0.115	0.142	0.03 (-1)
121	21331	0.113	0.139	0.03 (-)
122	31113	0.108	0.122	0.01 (+4)

Table S1: Full UK social EQ-5D 3L value sets for the CAM and the MUAP (continued).

Rank	Health state	CAM Value	MUAP Value	Difference Abs. (rank)	Rank	Health state	CAM Value	MUAP Value	Difference Abs. (rank)
123	32121	0.106	0.137	0.03 (-1)	184	33122	-0.084	-0.054	0.03 (-)
124	13313	0.104	0.117	0.01 (+4)	185	12133	-0.086	-0.061	0.03 (+1)
125	31222	0.103	0.130	0.03 (-2)	186	22332	-0.088	-0.054	0.03 (-3)
126	11332	0.101	0.129	0.03 (-2)	187	33321	-0.092	-0.067	0.02 (-)
127	22123	0.100	0.127	0.03 (-2)	188	23323	-0.097	-0.078	0.02 (+1)
128	33211	0.093	0.113	0.02 (+2)	189	32313	-0.102	-0.085	0.02 (+2)
129	23213	0.088	0.102	0.01 (+4)	190	23132	-0.102	-0.071	0.03 (-2)
130	22131	0.087	0.121	0.03 (-3)	191	23331	-0.110	-0.084	0.03 (-1)
131	21232	0.085	0.114	0.03 (-2)	192	31323	-0.114	-0.097	0.02 (+3)
132	12323	0.080	0.103	0.02 (-)	193	33113	-0.115	-0.102	0.01 (+4)
133	12132	0.075	0.110	0.04 (-2)	194	31132	-0.119	-0.090	0.03 (-2)
134	31213	0.071	0.083	0.01 (+6)	195	33222	-0.121	-0.093	0.03 (-2)
135	32221	0.069	0.099	0.03 (-1)	196	13332	-0.123	-0.095	0.03 (-2)
136	12331	0.067	0.097	0.03 (-1)	197	12233	-0.123	-0.099	0.02 (-1)
137	13123	0.066	0.087	0.02 (+1)	198	31331	-0.127	-0.103	0.02 (-)
138	23322	0.064	0.094	0.03 (-2)	199	21333	-0.132	-0.113	0.02 (+1)
139	22223	0.064	0.089	0.03 (-2)	200	23232	-0.139	-0.109	0.03 (-1)
140	32312	0.059	0.086	0.03 (-1)	201	32123	-0.140	-0.115	0.02 (-)
141	13131	0.053	0.081	0.03 (+1)	202	33213	-0.152	-0.140	0.01 (+3)
142	22231	0.051	0.083	0.03 (-1)	203	32131	-0.153	-0.121	0.03 (-1)
143	31322	0.047	0.075	0.03 (-)	204	31232	-0.155	-0.128	0.03 (-1)
144	33112	0.046	0.070	0.02 (+1)	205	22133	-0.158	-0.131	0.03 (-1)
145	12232	0.038	0.072	0.03 (-1)	206	33322	-0.176	-0.149	0.03 (-)
146	33311	0.038	0.057	0.02 (+1)	207	32223	-0.176	-0.153	0.02 (-)
147	23313	0.033	0.046	0.01 (+4)	208	12333	-0.178	-0.155	0.02 (-)
148	11133	0.032	0.051	0.02 (+1)	209	32231	-0.189	-0.160	0.03 (-)
149	13223	0.029	0.049	0.02 (+1)	210	13133	-0.192	-0.172	0.02 (+2)
150	21332	0.029	0.058	0.03 (-4)	211	23332	-0.194	-0.165	0.03 (-1)
151	32122	0.022	0.056	0.03 (-3)	212	22233	-0.194	-0.170	0.02 (-1)
152	13231	0.016	0.042	0.03 (+1)	213	33313	-0.207	-0.196	0.01 (+1)
153	31313	0.016	0.027	0.01 (+4)	214	31332	-0.211	-0.184	0.03 (-1)
154	32321	0.014	0.043	0.03 (-2)	215	13233	-0.229	-0.210	0.02 (+2)
155	33212	0.009	0.031	0.02 (+1)	216	32323	-0.232	-0.209	0.02 (-)
156	22323	0.008	0.033	0.02 (-1)	217	32132	-0.236	-0.202	0.03 (-2)
157	22132	0.003	0.040	0.04 (-3)	218	32331	-0.245	-0.215	0.03 (-)
158	33121	-0.000	0.027	0.03 (+1)	219	33123	-0.245	-0.226	0.02 (+1)
159	22331	-0.005	0.027	0.03 (-1)	220	22333	-0.250	-0.226	0.02 (-1)
160	11233	-0.005	0.013	0.02 (+3)	221	33131	-0.258	-0.232	0.03 (-)
161	23123	-0.005	0.017	0.02 (-)	222	23133	-0.263	-0.242	0.02 (+1)
162	32113	-0.010	0.009	0.02 (+3)	223	32232	-0.273	-0.241	0.03 (-1)
163	32222	-0.015	0.018	0.03 (-3)	224	31133	-0.280	-0.261	0.02 (-)
164	12332	-0.017	0.016	0.03 (-2)	225	33223	-0.282	-0.264	0.02 (-)
165	23131	-0.018	0.010	0.03 (-1)	226	13333	-0.284	-0.266	0.02 (-)
166	31123	-0.022	-0.003	0.02 (+2)	227	33231	-0.295	-0.270	0.02 (-)
167	13323	-0.026	-0.007	0.02 (+2)	228	23233	-0.300	-0.280	0.02 (-)
168	13132	-0.031	-0.000	0.03 (-1)	229	31233	-0.316	-0.299	0.02 (+1)
169	22232	-0.033	0.002	0.03 (-3)	230	32332	-0.328	-0.297	0.03 (-1)
170	31131	-0.035	-0.009	0.03 (-)	231	33323	-0.337	-0.320	0.02 (+1)
171	33221	-0.037	-0.012	0.03 (-)	232	33132	-0.342	-0.313	0.03 (-1)
172	13331	-0.039	-0.013	0.03 (-)	233	33331	-0.350	-0.326	0.02 (-)
173	21133	-0.040	-0.019	0.02 (-)	234	23333	-0.355	-0.336	0.02 (-)
174	23223	-0.042	-0.022	0.02 (-)	235	31333	-0.372	-0.355	0.02 (+1)
175	33312	-0.046	-0.024	0.02 (-)	236	33232	-0.379	-0.351	0.03 (-1)
176	32213	-0.046	-0.029	0.02 (+1)	237	32133	-0.398	-0.374	0.02 (-)
177	23231	-0.055	-0.028	0.03 (-1)	238	33332	-0.434	-0.407	0.03 (-)
178	31223	-0.058	-0.041	0.02 (+2)	239	32233	-0.434	-0.412	0.02 (-)
179	11333	-0.060	-0.043	0.02 (+2)	240	32333	-0.490	-0.468	0.02 (-)
180	13232	-0.068	-0.039	0.03 (-1)	241	33133	-0.503	-0.484	0.02 (-)
181	32322	-0.070	-0.038	0.03 (-3)	242	33233	-0.540	-0.522	0.02 (-)
182	31231	-0.071	-0.047	0.02 (-)	243	33333	-0.595	-0.578	0.02 (-)
183	21233	-0.077	-0.057	0.02 (+2)					