University of Southern Denmark · Syddansk Universitet

Modelling Danish EuroQol (EQ-5D) Tariffs by Applying the Time Trade-Off Method

Kim U. Wittrup-Jensen[1,4], Jørgen Lauridsen[2], Kjeld M. Pedersen[3]

(1): Bayer HealthCare AG, kim.wittrup-jensen@bayerhealthcare.com

(2): Institute of Public Health – Health Economics, E-mail jtl@sam.sdu.dk

(3): Institute of Public Health – Health Economics, University of Southern Denmark, kmp@sam.sdu.dk

(4): The study was done when Kim U. Wittrup-Jensen was a PhD student at University of Southern Denmark

Health EconomicsPapers
2008:4

UNIVERSITY OF SOUTHERN DENMARK

Abstract

Background: Notwithstanding the proposed use of cost-utility analysis (CUA) and Quality Adjusted Life-Years (QALYs) to inform health care priority setting in Denmark, to date there has been no research into Danes' preferences for health. For the time being QALYs implemented in Danish CUAs rely, at best, on preferences for health estimated in populations in other European countries e.g. the UK. However, Danes' preferences for health may not be equivalent to preferences in the UK. Hence estimation of QALYs and their application in Danish CUAs for priority purposes ought to be based upon Danes' preferences for health.

Objectives: To model a unique set of Danish EuroQol EQ-5D tariffs applying the Time Trade-Off (TTO) method. Second, to compare Danish EQ-5D tariffs with EQ-5D tariffs modelled in other countries. Third, to compare TTO elicited EQ-5D tariffs with Visual Analogue Scale (VAS)-based EQ-5D tariffs by estimating a power function describing the exponential and systematic relationship.

Data and methods: 1332 interviews were conducted in the respondents' own homes, where respondents were asked to assess hypothetical health states by the TTO method. In total, the EQ-5D classification system encompasses 243 health states, excluding the two states 'death' and 'unconscious'. However, as each respondent, at maximum, was capable of valuing around 13-30 health states, we applied a split-sample technique. 46 health states were directly valued by the respondents and used to estimate the remaining non-directly valued health states by applying different regression techniques and different model specifications.

Results: Based on the performance of multiple statistical tests, an additive model is chosen to be the most appropriate to estimate a national set of Danish EQ-5D tariffs. A cross-national comparison of country-specific EQ-5D tariffs, all estimated by the applying the TTO method and using the exact same model, reveals a high degree of correlation between the different EQ-5D sets. It appears that EQ-5D tariffs estimated in West European countries are very similar, but are quite different from Japanese EQ-5D tariffs. Finally, a power function describing the functional relationship between EQ-5D VAS and TTO valuations is estimated at the aggregate level, which turns out to be similar to previous findings.

Conclusions: The model appears to predict the values of the health states for which there are no direct observations. The implications are that the model can be used to interpolate values for health states where no direct observations exist. Finally, there are indications suggesting that instead of focusing on country-specific tariffs, the focus should be on modelling tariffs covering two countries or more, which are more or less alike regarding preferences for health.

Introduction

In 1993 the Measurement and Valuation of Health (MVH) Group at the Centre for Health Economics in York conducted an extensive interview-based study applying the Time Trade-Off (TTO) method, resulting in UK-based EuroQol (EQ-5D) tariffs [Dolan *et al.* 1995; Dolan *et al.* 1996b; Dolan 1997a; Badia *et al.* 2001]. Since then several countries around the world have replicated this study, however none are as large or representative as this study reported here [Badia *et al.* 1999; Fukada *et al.* 1999; Badia *et al.* 2001]. The main reason for estimating country-specific EQ-5D tariffs is that there is a strong belief, and evidence, that preference for health differs across country-specific borders [Badia *et al.* 2001].

In Denmark economic guidelines do not exist for the reimbursement of pharmaceuticals, however, it is advised that the cost-effectiveness of the pharmaceutical in question should be investigated. Currently Danish Cost-Utility Analyses (CUAs) and Health Technology Assessments (HTAs), using EQ-5D as the method for measuring outcome, will usually have to apply the UK-based EQ-5D tariffs in order to obtain a cardinal value for each of the 243 health states on a 0 (worst) to 1 (best) scale. As UK-based EQ-5D tariffs may not reflect Danes' preferences for health, there is a strong need for tariffs elicited directly within the general Danish population.

Objectives

The results presented here represent the first attempt to model and explore the health state preferences for EQ-5D tariffs of a randomised sample of the Danish general population. Specifically, the aim is to produce a tariff of health state preferences that will better inform economic evaluation in Denmark. A consensus emerging in the literature is that such a tariff should most appropriately be based on the valuations assigned to a set of hypothetical health states by the general public (in their dual roles as taxpayers and potential patients) behind 'a veil of ignorance'.

The health state classification system and preference-elicitation methods used in this study are outlined in the next section, followed by a discussion of methodological issues in the relationship between health description and health valuation, with an application to the EQ-5D classification system. In the following section, focus is on a description of the characteristics of the dataset. The subsequent section contains the model and regression techniques applied to analyse the data. Then the focus is upon how the Danish sets of tariffs relate and correlate with foreign – country-specific – EQ-5D tariffs. Next, a possible (exponential) relationship between VAS-based and TTO-based EQ-5D tariffs is explored by estimating a *power function*. In closing, attention is drawn to a number of remaining methodological issues for future research in this area.

Data and methods

The EQ-5D questionnaire

This study employs the health state classification system known as the EuroQol (EQ-5D), developed by the EuroQol Group, which is a consortium of investigators in Western Europe [Gold *et al.* 1996]. The EQ-5D profile classifies individuals into one of 243 possible health states, plus the states *dead* and *unconscious* [Brooks 1996]. Once the individual has put himself in such a health state, the researcher can assign a relevant value. Normally, empirical values are only available for a selection of all possible health states as the *respondent burden* imposed by valuation exercises limits the number of health states a respondent can sensibly be asked to value directly. However, by estimating a parametric relationship between the profile and the known values, it is possible to estimate the values for non-directly valued health states [Busschbach *et al.* 1999]. The idea put forward by the EuroQol Group is the assumption that health can be characterized by a set of scores applied to five aspects of health status. Normally such aspects are referred to as *dimensions*. Each of the five dimensions consists of three ordinal levels: (1) 'no problems', (2) 'some/moderate problems' and (3) 'extreme problems/unable to'. See Figure 1 for an illustration.

Figure 1. The EuroQol health dimensions and scores.

Dimensions	Levels	Scores
Mobility	No problems in walking about	1
	Some problems in walking about	2
	Confined to bed	3
Self-care	No problems with self-care	1
	Some problems with washing or dressing self	2
	Unable to wash or dress self	3
Usual activities	No problems with performing usual activities (e.g. work, study, housework, family or leisure activities)	1
	Some problems with performing usual activities	2
	Unable to perform usual activities	3
Pain/discomfort	No pain or discomfort	1
	Moderate pain or discomfort	2
	Extreme pain or discomfort	3
Anxiety/depression	Not anxious or depressed	1
	Moderately anxious or depressed	2
	Extremely anxious or depressed	3

Within the EQ-5D system every individual health state can be described by a row vector $\mathbf{x}(x_1, x_2, \ldots, x_5)$ in which the element x_i represents the score on the dimension *i*. This implies that x_1 = the score on 'mobility', x_2 = the score on 'self-care', x_3 = the score on 'usual activities', x_4 = the score on 'pain/discomfort' and x_5 = the score on 'anxiety/depression'. The score on a dimension is '1' if it is the highest level and '3' if it is the lowest (see also figure 1). For example, health state 21132 would indicate level 2 on mobility ('I have *some problems…*'), *no problems* with self-care or usual activities, *extreme* pain or discomfort and *moderate* anxiety or depression. In total, excluding the two health states *dead* and *unconcious*, the system encompasses 3^5 = 243 possible health states, even though some health states may seem unrealistic e.g. health state 31111 where the individual is *bedridden*, but has *no problems* with 'self-care' or 'usual activities'. In an ideal world all individuals ought to value all 243 health states, however, in practice they can only value around 13 to 30 health states each, which makes modelling unavoidable for the remaining health states [Dolan 1997a; Busschbach *et al.* 1999].

Randomisation of the data set

As the study aimed at estimating preferences for health reflecting those of the Danish population, it was important that the sample was randomised within the general Danish population. To what degree this was the case could not be measured directly. However, it could be measured indirectly by comparing socio-economic characteristics for the sample with the population as a whole, to check for significant differences. In the study (1,332 respondents) 18-29 year-olds were under-represented, while those aged 60 and above were over-represented.

Table 1. Randomisation judged by gender and age distributions. Per cent.

	General population (N = 4,127,847) (January 1st 2000)	**TTO personal interview survey (N = 1332) (Winter/Spring 2000)**
Gender		
Male	48.9 %	42.0 %
Female	51.1 %	58.0 %
Age		
18 – 29 years	21.1 %	15.8 %
30 – 59 years	54.1 %	55.1 %
≥ 60 years	24.8 %	29.1%

It was decided not to weigh the sample, even though a difference between females and males was present, which one should be aware of when performing analyses where gender plays a significant role. Weighting gender or age would not have had a significant impact on the results and, moreover, there could be adverse side effects such as an 'over-weighting' of cases with error in measurement or other inaccuracies. Weighting was a possibility, but for operational reasons the decision was made not to undertake such an exercise.

The data set

Face-to-face interviews were conducted during spring year 2000 with a randomised sample of the Danish population. A computer-assisted interviewing method was applied. Each interviewer had a lap-top computer and conducted the interview by reading from the screen (and showing props whenever relevant) and immediately keying in the response. This was not only easier than using paper questionnaires but also removed one of the sources of error, since paper-computer transfer was eliminated. In addition logical consistency checks were an integrated part of the computer set-up.

A total of 4,075 addresses were used. These were randomised within the Danish adult (i.e. ≥ 18 years with no upper age limit) population. 1,421 were not at home (after three contact attempts), leaving a net sample of 2,654 with whom contact was obtained. 1,322 refused to participate, resulting in a completion rate of 50 per cent, which is fairly low for face-to-face interviews in Denmark. A total of 1,332 completed interviews were included in the study.

The study was interview-based and was split into four different exercises: 1) the EQ-5D profile questionnaire, 2) the ranking exercise, 3) the valuation exercise and 4) the TTO exercise. Finally, each respondent filled out background questions. The exercises (1) – (3) were so-called 'warm-up' exercises, where the idea was to make respondents familiar with each health state and the *way of thinking* concerning how to value the EQ-5D health states within a TTO context.

Split-samples

Pilot surveys have shown that no respondents are capable of valuing more than 13–30 health states within the same exercise [Dolan *et al.* 1995]. However, this number was not enough for a direct valuation of all 243 possible EQ-5D health states. Hence it was necessary to estimate the remaining non-directly valued health states, based on a modelling exercise of the directly valued health states. By obtaining as many directly-valued health states as possible the prediction of the remaining non directly-valued health states became more accurate. Thus the survey was structured as a split-sample study, where respondents were split into four equally-sized groups, where each group was presented with different health states which the ranked and valued.

In choosing which health states were to be included in the study it was important that they covered a representative sample of all 243 health states. Consequently, it was important to include as many different combinations across the different dimensions as possible. Furthermore, it was imperative that the health states were plausible for the respondents. For example, level 1 (no problems) in the dimension 'usual activities' was not combined with level 3 (confined to bed) in the dimension 'mobility'. Figure 2 gives an overview of those health states that were chosen for direct valuation, while Table 2 is an overview of the health states included in each of the four split-samples.

Figure 2. EQ-5D health states valued directly in the study.

i)	*All respondents valued:* - EQ-5D health states 11111, 22222, 33333 and 'dead' - 2 out of 5 'mild' EQ-5D health states: 21111, 12111, 11211, 11121, 11112 - 8 other EQ-5D health states (used in earlier EuroQol investigations) which covered both less serious and more serious health states - 2 relevant EQ-5D health states (which Danish diabetes and/or heart disease patients had used to describe their own health) to be used to test VAS and TTO tariffs. Hence each respondent valued 16 EQ-5D health states.
ii)	All 4 samples included states that covered levels 1,2,3 of all 5 dimensions in EQ-5D
iii)	All EuroQol 'core ' states were included

Table 2. Distribution of health states across the four split-samples.

Split sample 1	Split sample 2	Split sample 3	Split sample 4	Commentary
n=340	n=335	n=333	n=324	
11111	11111	11111	11111	*The four common states in all four variants*
22222	22222	22222	22222	
33333	33333	33333	33333	
Dead	Dead	Dead	Dead	
21111	11112	11211	21111	*2 of the 5 'mild' States*
11211	12111	11121	11112	
12211	11131	23313	32232	*Eight states used in earlier EQ-5D Investigations covering all levels*
11113	22121	11122	11312	
21222	21312	13311	11133	
32331	12222	22122	33321	
23321	32211	22331	21323	
21232	12223	22112	13212	
23232	21133	32313	Unconscious	
22233	22323	32223	33232	
11212	21221	21121	11221	*Two comparison states with other investiga-tions*
22333	33322	21322	22322	

The Time Trade-Off (TTO) exercise

In the TTO exercise a specially designed board was applied to illustrate the trade-offs between the EQ-5D health states. One side of the board was of relevance for respondents who valued health states as better than death and the other side was for respondents who valued health states as worse than death. In the former case a so-called 'equivalence technique' was applied, where the respondent was asked to assess how long a time (x) spent in health state 11111 he or she thought was equivalent to spending 10 years in a specific health state; the shorter the time the respondent was willing to spend in the health state, the worse the health state. Respondents also had the option not to trade-off any time. In cases where a health state was assessed to be worse than death, the trade-off was between dying immediately or spending a given amount of time ($10 - x$) in the specific health state, followed by x years in the health state 11111; the more time necessary in health state 11111 in order to compensate for the less time in the specific health state, the worse the health state.

If 'perfect health' and 'death' are given the values 1 and 0, respectively, then the values for health states valued better than death in the TTO exercise were given by the formula $x/10$, where x was the number of years spent in full health. For health states worse than death, the values were given by the formula $-x/(10\text{-}x)$. Consequently negative values were calculated on a ratio scale (not an interval scale), and by contrast with health states assessed to be better than death, these health states were in theory without a lower limit. In agreement with the similar Measurement and Valuation of Health (MVH) study performed in the UK, the lower limit was explicitly fixed at -39 [Dolan 1997a]. Problems that may occur as a result of this asymmetry between positive and negative values are dealt with in the discussion section.

The data set used in the regression modelling to predict non-directly valued EQ-5D health states includes all logical inconsistencies (n = 1,332). A logical inconsistency could, for example, be a respondent valuing the EQ-5D health state 13111 as better (i.e. higher value) than the EQ-5D health state 12111 on a cardinal scale from 0 (dead) to 1 (perfect health). The best and probably the right way to handle inconsistencies would be to exclude these from further analysis. However, exclusion of all (strong) inconsistencies reduced the present data set to cover valuations from around 29 respondents (around 2 per cent of the original data set including all (strong) inconsistencies). Consequently, it was chosen to include all (strong) inconsistencies in the modelling.

A regression model was applied where the functional form explicitly was assumed to be additive.[1] The dependent variable was defined as $1 - S$, where S was the value for a given health sta-te. In addition to applying an intercept, the independent variables were defined based on the EQ-5D ordinal structure. In total, three sets of dummy variables were created:

[1] See the dicussion on this topic in Busschbach et al. (1999).

1. Two dummy variables for every dimension, one that represented movement between the three levels, and one that represented movement from level 2 to level 3. This allowed the effect of moving from level 1 to level 2 to be different from the effect of moving from level 2 to level 3.
2. Dummy variables allowing (first-order) interaction between the five dimensions.
3. Dummy variables that captured how many times a health state contained one or more dimensions on level 1 or level 3.

Figure 3 gives an overview of dummy variables and their definitions. Furthermore, two variables were specified: N2 and N3. These two variables should capture whether one of the levels within the five dimensions was on level 2 or 3, respectively.

In the TTO exercise the asymmetry between positive and negative values caused problems, since those respondents who valued a health state as being worse than death would have a larger impact on the predicted values in the model than those respondents who valued the same health state as being better than death. According to Patrick *et al.* (1999) the negative values can be transformed in such a manner that the score for health states valued worse than death is explicitly truncated to -1. The justification for this transformation is statistical, but it may also be justified from a purely psychometric point of view. Perhaps respondents assess the scale for health states being worse than death in the same manner as they ought to assess the scale for health states being better than death, i.e. as an interval (not a ratio) scale. In this case health states worse than death were transformed using the formula $(x/10) - 1$, where x represents the number of years spent in the best health state (11111).

The bulk of the extensive programming was performed in the statistical software program *SAS* 6.2e, but at some stages the statistical software programme *STATA* was applied [SAS Institute Staff 2000; Rabe-Hesketh *et al.* 2000].

Figure 3. Definition of variables used in the regression analysis.

Variable	Definition
a	Intercept: indicator for any movement away from 11111 (perfect health)
MO	1 if mobility is level 2; 2 if mobility is level 3; 0 otherwise
SC	1 if self-care is level 2; 2 if self-care is level 3; 0 otherwise
UA	1 if usual activities is level 2; 2 if usual activities is level 3; 0 otherwise
PD	1 if pain/discomfort is level 2; 2 if pain/discomfort is level 3; 0 otherwise
AD	1 if anxiety/depression is level 2; 2 if anxiety/depression is level 3; 0 otherwise
M2	1 if mobility is level 3; 0 otherwise
S2	1 if self-care is level 3; 0 otherwise
U2	1 if usual activities is level 3; 0 otherwise
P2	1 if pain/discomfort is level 3; 0 otherwise
A2	1 if anxiety/depression is level 3; 0 otherwise
MOSC	The product of MO and SC
MOSC	The product of MO and UA
MOPD	The product of MO and PD
MOAD	The product of MO and AD
SCUA	The product of SC and UA
SCPD	The product of SC and UA
SCAD	The product of SC and UA
UAPD	The product of UA and PD
UAAD	The product of UA and AD
PDAD	The product of PD and AD
F11	1 if the health state contains 1 dimension at level 1; otherwise 0
F21	1 if the health state contains 2 dimensions at level 1; otherwise 0
F31	1 if the health state contains 3 dimensions at level 1; otherwise 0
F41	1 if the health state contains 4 dimensions at level 1; otherwise 0
F13	1 if the health state contains 1 dimension at level 3; otherwise 0
F23	1 if the health state contains 2 dimensions at level 3; otherwise 0
F33	1 if the health state contains 3 dimensions at level 3; otherwise 0
F43	1 if the health state contains 4 dimensions at level 3; otherwise 0
F53	1 if the health state contains 5 dimensions at level 3; otherwise 0
N2	1 if any dimension is at level 2; otherwise 0
N3	1 if any dimension is at level 3; otherwise 0

The analysis was conducted at the individual level, where each individual valued a discrete number of health states. In such a situation it is important to bear in mind that there obviously had to be a particular connection between the valuations of the health states. In other words, if a re-spondent values a health state which lies below the mean compared to the valuations given by the rest of the sample, there is a tendency that the respondent will also value the remaining health states below the sample mean. This implies that the variance of the error term is partly determined by those respondents who value the health states. However, this violates one of the assumptions for using the *Ordinary Least-*

Squares (OLS) model, so this method could not be applied here. Instead a *Random Effects (RE)* model, which was specified as follows, was applied:

$$y_{it} = \alpha_1 + \beta_2 x_{2,it} + \beta_3 x_{3,it} + \ldots + \beta_K x_{K,it} + u_i + \varepsilon_{it} \qquad (1)$$

where u_i was an individual-specific random term representing the extent to which the intercept of the *i*'th respondent differs from the overall intercept, α_1. Rewriting (1) as

$$y_{it} = (\alpha_1 + u_{1,i}) + \beta_2 x_{2,it} + \beta_3 x_{3,it} + \ldots + \beta_K x_{K,it} + \varepsilon_{it} \qquad (2)$$

or

$$y_{it} = \beta_1 + \beta_2 x_{2,it} + \beta_3 x_{3,it} + \ldots + \beta_K x_{K,it} + \varepsilon_{it} \qquad (3)$$

we may equivalently view the RE model as a *Random Intercept* model. The randomisation of the intercept β_1 may equally well be applied to any of the parameters by defining

$$\beta_K = \alpha_K + u_{k,i} \qquad (4)$$

where $u_{k,i}$ is the individual specific error term representing the extent to which the coefficient for x_k of the *i*'th respondent differed from the average fixed coefficient α_k. As shown, this *random coefficient (RC)* model was a general formulation covering the RE model as a special case in which only the intercept was randomised.

The RE model may be tested against the RC model using a split-sample test. If the RE- and RC-coefficients were not significantly different, then the RE model was more efficient than the RC model, whereas significant differences indicate efficiency problems in the RE model. The split-sample test was performed by randomly splitting the sample of respondents into two equally sized sub-samples, estimating the RC model on sub-sample 1 and the RE model on sub-sample 2. Next, the test size $\mathbf{d}'\mathbf{V}^{-1}\mathbf{d}$ where **d** measured the distance between the RE and RC parameter vectors and *V* the sum of their covariance matrix, was calculated. Finally, the test size was compared to a χ^2 distribution with *df* equal to *K*.

Further, the efficiency of the RE model versus the OLS model was considered. If the RE model did not represent any improvement over the OLS model, then the RE model was less efficient than the OLS model. On the other hand, if the RE model represented an improvement, then the RE model was more efficient. Thus, the efficiency of the OLS model needed to be tested. Three tests were performed. The first was a *Lagrange Multiplier (LM)* test, based on the OLS estimates only. Due to the conservative nature of the LM test (i.e. a relatively large tendency to reject the OLS model in finite-sized samples), a *Likelihood Ratio (LR)* test based on the difference between the OLS model and the RE model log likelihood values, and a *Wald* test based on the RE model results only (calculated as the squared *t*-value for significance of the estimated variance of u_i) was performed.

An important consistency issue is frequently ignored in empirical work. A condition for consistency of the RE model is that the individual effects are uncorrelated with the explanatory variables. In the present case this implied that the individual deviations from the average valuations of the health states did not vary over the range of health states. Thus, if a respondent overestimated each of the health states presented by a certain amount, then this amount should not change if the respondent was presented with any other health state. Clearly, this property is not automatically guaranteed. In the case of violation of the independence assumption, the *Fixed Effect (FE)* model would provide consistent estimates. In the FE model specification, the random effects were replaced with a fixed effect for each respondent. Thus, the FE model was similar to an OLS with a dummy variable for each respondent. On the other hand, if the RE model was consistent, then we would have expected the FE model to be strongly inefficient because many of the intercept parameters - one per respondent - may be without significant difference. In order to test the consistency of the RE model, we calculated the *Hausman* test for equality of the RE and FE parameters.

The models were tested for misspecification using a *Ramsey RESET* test and a test for general *heteroscedasticity*. Both tests are two-stage tests with a common first stage consisting of estimating the model in question. For the RESET, the second step consists of re-estimating the model with the squared predicted values from the first stage as an additional explanatory variable, and using the significance of the *F-value* of this variable as indicative of *functional misspecification*. For the heteroscedasticity test, the second stage consisted of regressing the predicted values from the first stage on the squared residuals from the first stage, and using the significance of the *F*-value for this squared residual as indicative of heteroscedasticity.

In order to measure the extent of *multicollinearity*, the *condition number* (CN) for each model was calculated and used the standard 'rule-of-thumb' that a CN in excess of 20-30 indicates a problem.

Finally, to make a choice between different ways of representing a relationship between valuations and health states, it was required that the predicted valuations should be logically consistent. For a model with a specific set of explanatory variables – if it was possible while maintaining the requested logical consistency – both efficiency and consistency were required. This choice was based on the test statistics for the RE, FE, OLS and RC models versus each other, as described above. In case of conflicts between the logical consistency request and the test statistics, it was decided to choose the logically consistent model for further consideration, while reporting and discussing the statistically optimal model in order to shed light on possible problems in the chosen model.

It turned out that the directly valued TTO health states (n=46) were strongly skewed to the right (i.e. many health states were valued around +1). In order to compensate for this, a *censored Tobit* model was estimated using *STATA*, as this option does not exist in *SAS*. As noted by Austin *et al.* (2000), regression methods that ignore the presence of censoring in the health status measurements can produce biased coefficient estimates. Also the ML estimation of the RE model was estimated using the software program *STATA*, to detect any differences compared to *SAS*. Finally, a *Generalized Least Square* (GLS) model matching the model used in Dolan (1997) was estimated.

Results

The EQ-5D profile questionnaire

All respondents were asked to fill out the EQ-5D profile questionnaire, which results in an ordinal score. All those respondents who completed the interview (n=1,332), filled-out the EQ-5D profile questionnaire. As illustrated in Table 3, respondents placed themselves in 48 of the 243 possible health states. Over 60 per cent of the respondents had, according to themselves, perfect health (symbolised as 11111), while around 12 per cent of the remaining 40 per cent, had moderate 'pain/discomfort'. Furthermore, around 4 per cent also had problems with 'usual activities'.

Table 3. Distribution of EQ-5D health states based on results from the EQ-5D profile.
(n = 1,332).

Health state	Number of respondents	Per cent	Health state	Number of respondents	Per cent
11111	799	60.0	21211	9	0.7
11112	44	3.3	21212	1	0.1
11113	4	0.3	21221	38	2.9
11121	165	12.4	21222	19	1.4
11122	26	2.0	21231	9	0.7
11131	2	0.2	21232	4	0.3
11211	16	1.2	21311	1	0.1
11212	3	0.2	21321	7	0.5
11221	50	3.8	21322	2	0.2
11222	9	0.7	21331	5	0.4
11231	3	0.2	21332	2	0.2
11232	1	0.1	22121	3	0.2
11311	1	0.1	22122	1	0.1
11321	1	0.1	22133	1	0.1
11331	1	0.1	22221	8	0.6
12111	1	0.1	22222	4	0.3
12221	4	0.3	22231	4	0.3
12222	3	0.2	22232	2	0.2
21111	15	1.1	22312	1	0.1
21112	2	0.2	22321	3	0.2
21121	40	3.0	22322	1	0.1

21122	7	0.5	22331	3	0.2
21131	4	0.3	23321	1	0.1
21132	1	0.1	31332	1	0.1

The ranking exercise

In the rank ordering exercise the respondents had all health states (n = 16 for each of the four samples) placed in front of them and were asked to rank them so that the best health state was on *top* and the worst health state at the *bottom*. The respondents were told that they had to picture themselves in each health state for a period that would last *10 years* after which they would die.

Table 4 shows the respondents' mutual ranking of the 16 health states. As seen from the table, judged by median value, respondents in all four split-sample surveys ranked the health state 11111 (perfect health) as the best health state, which follows the method in the EQ-5D classification system. Furthermore, the table illustrates that the worse the dimension the lower the health state was ranked, compared to the remaining health states. The health states 'Die Immediately' (DI) and 33333 were, for all four split-samples, placed at the bottom, while the health state 'UNConscious (UNC)' in split-sample III was assessed to be just as good/bad as DI.

Table 4. Rank ordering of health states judged by median value.

Rank	Split sample I (n = 340)	Split sample II (n = 335)	Split sample III (n = 333)	Split sample IV (n = 324)
1	11111	11111	11111	11111
2	11211, 21111	12111	11211, 11121	11112, 21111
3		11112		
4	12211	22121, 21221	21121, 11122	11221
5	11212			22222
6	21222	12222	22112	11312
7	22222	22222	22122	13212
8	11113	21312	22222	22322
9	23321, 21232	12223, 11131	21322, 13311	11133, 21323
10		32211		
11	22233, 23232		22331	32232
12		22323, 21133	32313, 23313	33232, 33321
13	32331, 22333	33322	32223	
14				DI, UNC, 33333
15	DI, 33333	DI, 33333	DI, 33333	
16				

Note: DI = Die Immediately and UNC = UNConscious.

The valuation exercise

In the valuation exercise the respondents were asked to value the same EQ-5D health states as in the ranking exercise, with the same time frame, on a scale ranging from 0 (worst) to 100 (best). This is also referred to as the Visual Analogue Scale (VAS) exercise. The results are illustrated in Table 5, where it can be seen that the valuations of the EQ-5D health states followed the same pattern as in the ranking exercise; the more serious the health state, the lower the numeric value. In all four split-sample surveys the health state 11111 was assessed at a value close to 100. At the other end of the scale, for all four split-samples, the health state 33333 was assessed to have a value close to 0. It is interesting that the health state DI in all four split-samples was assessed to be better than the health state 33333, indicating this health state to be worse than death. This was also the case for the health state 33322 in split-sample I, and the state UNC in split-sample IV.

Table 5. Mean value for EQ-5D health states in the VAS valuation exercise. (Not rescaled).

Split sample I (n = 340)		Split sample II (n = 335)		Split sample III (n = 333)		Split sample IV (n = 324)	
Health state	Value	Health state	Value	Health State	Value	Health state	Value
11111	98.5	11111	98.7	11111	99.3	11111	99.0
23321	37.5	32211	30.4	21121	76.3	11112	84.8
11211	86.2	11112	78.9	11211	86.3	DI	13.9
DI	13.8	DI	15.4	DI	12.1	22322	45.9
11212	71.4	12222	57.6	21322	44.1	11221	78.0
21232	37.7	21312	47.3	32223	17.0	33232	16.1
12211	74.3	33322	13.7	22331	28.1	32232	21.6
11113	42.3	22323	21.9	13311	40.3	11133	30.7
32331	17.5	12223	32.2	32313	15.9	33321	20.1
22333	15.7	11131	40.9	22222	45.9	21323	33.0
22233	20.8	12111	84.5	22112	62.0	13212	48.7
22222	49.8	22222	47.4	11121	84.5	22222	58.0
21111	84.6	22121	65.1	22122	57.8	UNC	11.3
21222	57.8	33333	3.7	23313	20.2	21111	86.7
33333	4.5	21133	21.8	33333	3.7	33333	4.3
23232	25.2	21221	63.7	11122	71.1	11312	60.6

Note: DI = Die Immediately, UNC = UNConscious.

The TTO exercise

By carefully checking the existing literature, eleven different model specifications were selected [Dolan 1997a; Rupel & Rebolj 2000; Devlin *et al.* 2000]. Overviews of these specifications are shown in Figure 4.

Figure 4. Overview of the models tested in predicting Danish EQ-5D tariffs.

Name = f(x)
TTO1 = f(MO, SC, UA, PD, AD)
TTO2 = f(MO, SC, UA, PD, AD, N3)
TTO3 = f(MO, SC, UA, PD, AD, M2, S2, U2, P2, A2)
TTO4 = f(MO, SC, UA, PD, AD, M2, S2, U2, P2, A2, N3)
TTO5 = f(MO, SC, UA, PD, AD, M2, S2, U2, P2, A2, MOSC, MOUA, MOPD, MOAD, SCUA, SCPD, SCAD, UAPD, UAAD, PDAD)
TTO6 = f(MO, SC, UA, PD, AD, MOSC, MOUA, MOPD, MOAD, SCUA, SCPD, SCAD, UAPD, UAAD, PDAD)
TTO7 = f(F11, F21, F31, F41, F13, F23, F33, F43, F53)
TTO8 = f(MO, SC, UA, PD, AD, F11, F21, F31, F41, F13, F23, F33, F43, F53)
TTO9 = f(MO, SC, UA, PD, AD, F13, F23, F33, F43, F53)
TTO10 = f(MO, SC, UA, PD, AD, M2, S2, U2, P2, A2, N2, N3)
TTO11 = f(MO, SC, UA, PD, AD, N2, N3)

First the RC model (R^2=0.41) was estimated, which resulted in inconsistent parameters, as some were negative. The split-sample test showed a strong significant difference between the RC- and RE-parameters, indicating efficiency problems in the RE model. However, since the parameters in the RC model were logically inconsistent, the RE model was chosen.

The LM test led to a rejection of the OLS model, indicating that the OLS model was inefficient. Furthermore, the LR test indicated that the OLS model should be rejected compared with the RE model. Finally, the Wald test showed that the RE model was efficient and implicitly rejected the OLS model. The ordering between the three tests is normally that Wald<LR<LM, implying that the LM test rejects OLS 'more easily' than both the LR and the Wald tests. The Hausman test showed that both Prob(H_0: REM)≈0 and Prob(H_0: OLS)≈0, which means that both the RE model and the OLS model are rejected. The implications are that the FE model performed better, with regard to significance, than both the RE and OLS models. However, the FE model produced inconsistent parameters, as some parameters were negative. Therefore the RE model was chosen.

The RE model failed the Ramsey RESET test ($p<0.0001$) indicating that the model suffered from misspecification. In addition, the model also suffered from heteroscedasticity ($p<0.0001$). The RE model did not, however, suffer from multicollinearity since the CN was 12.21.

The modelling in *STATA* showed that the ML estimation of the RE model matched the model estimated in *SAS*. The Tobit censored RE model resulted in inconsistent parameters and, finally, the GLS model was, as expected, a perfect match of the ML results obtained from the RE model in *SAS*.

From a statistical point of view the FE model was superior to the RE model. However, some of the parameters in the FE model were negative. Hence, in estimating Danish EQ-5D tariffs, the RE model was applied. The detection of both misspecification, i.e. incorrect functional form & (omitted) variables, and heteroscedasticity in the RE model is not surprising, given that the power of the RESET tests increased as the sample size increased. The presence of heteroscedasticity is likely to be one of the causes of inefficient parameter estimates. Despite trying to compensate for the direct valuations being highly skewed, the censored Tobit model resulted in inconsistent parameters. However, a comparison of the censored Tobit model with the uncensored GLS model, shows that the coefficients for MO – AD and N3 (for model TTO4) were lower in the GLS model and that the coefficients for M2 – A2 were higher in the GLS model. Because of this problem with skewed data the Tobit model is believed to be the most accurate. The GLS model ignores the censoring of the 'good' health states (i.e. cards with an over-weigh-ting of 1's and 2's) and puts too much emphasis on the decline from 2 to 3. Finally, by rejecting the FE model, possibility that individual socio-economic characteristics could have a significant influence on how people valued health was implicitly rejected. In other words, factors such as social class, gender, and age did not seem to influence the way people valued health. However, this is a rather strong postulation and more research is needed to say anything conclusive.

First, all eleven models were regressed using the SAS statistical software package. As expected all models yielded a high R^2 value. The models TTO5, TTO6 and TTO11 all resulted in both inconsistent (wrong sign) and insignificant parameters at the 1 per cent level. The models TTO1, TTO2, TTO7, TTO8, TTO9, and TTO10 resulted in inconsistent parameters, but yielded significant parameters at the 1 per cent level. TTO4 had consistent parameters, but an insignificant parameter for the dimension of 'usual activities' at the 10 per cent level. Finally, the TTO3 model resulted in both consistent and significant parameters at the 1 per cent level. The intermediate result was that the models that allowed for interaction between the five dimensions did not improve the results and almost all models had inconsistent parameters (some negative). From a purely objective point of view, the TTO3 model is the best model, i.e. *best fitting*, since the *p*-values for all parameters were significant ($p < 0.001$) and all parameters were consistent. However, since one of our aims was to compare our results with the UK-based tariffs, the estimates from TTO4 were used. The estimated parameters, *p*-values, and R^2 values for TTO3 and TTO4 are illustrated in Table 6.

Table 6. Estimates of model TTO3 and TTO4. *P*-values in brackets. (n = 1,332).

Estimates	**TTO3**	**TTO4**
Intercept (a)	0.1137 (<0.0001)	0.0881 (<0.0001)
MO	0.0532 (<0.0001)	0.0554 (<0.0001)
SC	0.0629 (<0.0001)	0.0658 (<0.0001)
UA	0.0478 (<0.0001)	0.0223 (0.0095)
PD	0.0623 (<0.0001)	0.0764 (<0.0001)
AD	0.0682 (<0.0001)	0.0594 (<0.0001)
M2	0.3048 (<0.0001)	0.2949 (<0.0001)
S2	0.0659 (<0.0001)	0.0470 (0.0005)
U2	0.0480 (0.0006)	0.0106 (0.4505)*
P2	0.2717 (<0.0001)	0.1926 (<0.0001)
A2	0.2314 (<0.0001)	0.2007 (<0.0001)
N3	-	0.1592 (<0.0001)
R^2	0.6551	0.6602

*(p > 0.10).

As noted by Devlin *et al.* (2000), the intercept (constant) applied in the modelling ought to be zero because when the dummies are set to zero (corresponding to the health state 11111, valued at unity) then $1 - S$ = intercept, which can be rearranged to $S = 1$ – intercept. The intercept in the model selected in this study (i.e. TTO3 and TTO4) was positive, consistent with the findings reported by Devlin *et al.* (2000) and Dolan (1997a). According to Dolan (1997a) the constant represents any move away from full health. Because of this the intercept represents a discontinuity in the model between level 1 and 2 in much the same way as the N3 term represents a discontinuity between level 2 and 3.[2] This interpretation was accepted and consequently a set of tariffs based on the sample of 1,332 individuals are estimated.

After testing eleven different models, the TTO3 model fitted the data very well (in terms of *goodness-of-fit* statistics) and this is readily interpretable as a main effects model, in which each of the five dimensions is independent of the others. Thus, the regression equation is expressed as follows:

$$Y = \alpha + \beta_1 MO + \beta_2 SC + \beta_3 UA + \beta_4 PD + \beta_5 AD + \beta_6 M2 + \beta_7 S2 + \beta_8 U2 + \beta_9 P2 + \beta_{10} A2 \quad (5)$$

where the TTO scores were explained by eleven independent variables: two variables for each dimension (one to represent the move from level 1 to 2 and one to represent the move from level 2 to 3), and an intercept (the interpretation of which was discussed earlier). Table 6 shows that the intercept was highly significant, suggesting that any move away from full health was associated with a substantial loss of utility. The largest decrement for a move from level 1 to level 2 was associated with 'anxiety/depression'. However, the corresponding move on the other four dimensions was very similar. 'Mobility' level 3 (confined to bed) dominated the weighting for level 3, some five to six times greater than that for the corresponding move on the 'usual activities' and 'self-care' dimensions, respectively. For the 'mobility', 'pain/discomfort' and 'anxiety/depression' dimensions, the move from level 2 to level 3 was seen to involve a much greater decrement than the move from level 1 to level 2.

As an example of how the tariff is generated, consider the EQ-5D health state 12233, which is estimated based on the TTO3 model. The result is illustrated in Figure 5.

Figure 5. How the preference-based EQ-5D tariffs are generated (example health state 12233).

Full health	=	1.0000
Intercept term (for any dysfunctional health state):		-0.1137
Mobility: level 1		-0.0000
Self-care: level 2 (1 x SC)		-0.0629
Usual activities: level 2 (1 x UA)		-0.0478
Pain/depression: level 3 (2 x PD + 1 x P2)		-0.3963
Anxiety/discomfort: level 3 (2 x AD + 1 x A2)		-0.3678
Therefore, the estimated value for EQ-5D health state 12233	=	0.0120*

Note: Based on results from the TTO3 model.

A way to validate the TTO tariffs is to compare the predicted values with the observed values for the 46 directly valued health states and see whether there are any major differences. The result of this comparison is shown in Table 7. As with the predicted valuations, the 46 observed valuations for the EQ-5D health states have been rescaled to fit the interval –1.0 to +1.0.[3]

[2] What the N2 should be representing is considered in the discussion section.
[3] As some valuations after the rescaling still yielded scores below –1.0, we truncated these values to –1.0.

Table 7. Predicted versus obs. valuations for directly valued health states in the TTO exercise.

Health states (n = 46)	Appears in sample number	Number of respondents	Predicted Value	Observed Value	Difference
23321	1	340	0.436	0.322	0.114
11211	1,3	673	0.838	0.877	-0.039
11212	1	340	0.770	0.786	-0.016
21232	1	340	0.321	0.195	0.126
12211	1	340	0.776	0.811	-0.035
11113	1	340	0.519	0.250	0.269
32331	1	340	-0.127	-0.307	0.180
22333	1	340	-0.137	-0.268	0.131
22233	1	340	-0.041	-0.148	0.107
22222	1,2,3,4	1332	0.592	0.584	0.008
21111	1,4	664	0.833	0.881	-0.048
21222	1	340	0.655	0.596	0.059
33333	1,2,3,4	1332	-0.624	-0.649	0.025
23232	1	340	0.129	-0.039	0.168
32211	2	335	0.364	0.196	0.168
11112	2,4	659	0.818	0.816	0.002
12222	2	335	0.645	0.649	-0.004
21312	2	335	0.621	0.628	-0.007
33322	2	335	0.009	-0.186	0.195
22323	2	335	0.197	0.155	0.042
12223	2	335	0.346	0.270	0.076
11131	2	335	0.490	0.288	0.202
12111	2	335	0.823	0.902	-0.079
22121	2	335	0.708	0.760	-0.052
21133	2	335	0.069	-0.020	0.089
21221	2	335	0.723	0.765	-0.042
21121	3	333	0.771	0.817	-0.046
21322	3	333	0.559	0.524	0.035
32223	3	333	-0.066	-0.238	0.172
22331	3	333	0.231	-0.060	0.291
13311	3	333	0.551	0.510	0.041
32313	3	333	-0.099	-0.235	0.136
22112	3	333	0.702	0.713	-0.011
11121	3	333	0.824	0.867	-0.043
22122	3	333	0.640	0.668	-0.028
23313	3	333	0.130	-0.083	0.213
11122	3	333	0.756	0.770	-0.014
22322	4	324	0.496	0.538	-0.042
11221	4	324	0.776	0.837	-0.061
33232	4	324	-0.229	-0.359	0.130
32232	4	324	-0.100	-0.186	0.086
11133	4	324	0.123	0.018	0.105
33321	4	324	0.078	-0.120	0.198
21323	4	324	0.260	0.196	0.064
13212	4	324	0.579	0.588	-0.009
11312	4	324	0.674	0.742	-0.068

The (absolute) mean difference between predicted and observed values for the directly valued health states was 0.061. In the MVH study the corresponding figure was 0.039 [Dolan 1997].[4] With regard to both the absolute and the relative mean differences between predicted and observed valuations, the UK valuations appear to be lower than the Danish valuations.

Comparison between country-specific TTO-based EQ-5D tariffs

One of the objectives of this study was to compare Danish EQ-5D tariffs with EQ-5D tariffs - all elicited by the TTO method - in other countries, to investigate whether preferences for health vary across countries. Indeed, one of the original aims of the EuroQol Group was to compare preference values for health states across countries [Brooks 1996]. For a variety of reasons, published studies of this kind are rare. However, a recent example is a comparison of UK and Spanish EQ-5D based TTO tariffs [Badia 2001].

There are no firmly established 'rules' or methodology for such comparisons. Two obvious pre-conditions are the use of roughly the same methodology and roughly similar respondents, i.e. (national) samples from well-defined populations and (roughly) the same socio-economic composition (to the extent that it is assumed that this may influence EQ-5D tariffs). Ideally, one should use a pooled individual-level data set to estimate the non-directly valued health states by means of the same methodology (regression method) and prediction equation. This appears to have been the case in the UK-.Spanish comparison. However, this is not (yet) possible for the Danish case. Hence this study proceeded in a different manner – using the published country-specific parameters for estimating national EQ-5D tariffs. National EQ-5D parameters from the UK [Dolan 1997], Japan [Fukuda *et al.* 1999], Spain [Badia *et al.* 2001], and, finally, Danish EQ-5D tariffs, all based on the TTO4 model, are presented in Table 8.

[4] The absolute mean difference is presented here, since the objective is to compare our results with the corresponding results found by Dolan (1997). However it can make a substantial difference whether one focuses on either the absolute or the relative mean difference between the predicted and observed values. Only reporting the absolute difference can give a false impression of how well the two observations relate, due to negative differences eliminating positive differences (or vice versa). In stead, reporting the relative mean difference gives a 'truer' picture of how well the two values really coincide. Hence, the relative mean difference between the predicted and observed values was 0.087. The corresponding relative mean difference in the UK MVH study was 0.053, however, this was not reported in Dolan [1997].

Table 8. TTO parameters for Denmark, the UK, Japan, and Spain used to estimate country-specific EQ-5D tariffs. All based on the TTO4 model.

Dimension	**Denmark**	**UK***	**Japan****	**Spain**
Constant	0.088	0.081	0.148	0.024
Mobility				
Level 2	0.055	0.069	0.078	0.106
Level 3	0.405	0.314	0.418	0.430
Self-care				
Level 2	0.066	0.104	0.053	0.134
Level 3	0.179	0.214	0.101	0.309
Usual activity				
Level 2	0.022	0.036	0.040	0.071
Level 3	0.055	0.094	0.128	0.195
Pain/discomfort				
Level 2	0.076	0.123	0.083	0.089
Level 3	0.345	0.386	0.189	0.261
Anxiety/depression				
Level 2	0.059	0.071	0.062	0.062
Level 3	0.319	0.236	0.108	0.144
N3	0.159	0.269	0.014	0.291
R^2	0.66	0.46	0.40	0.60

Source: *Badia *et al.* (2001) & **Tsuchiya A. *et al.* (2002).

Assuming that the pre-conditions noted above are present (as they mostly are for at least the UK, Spanish and Danish cases) one could ask: a) are the national rank orders of the health states similar, i.e. at the ordinal level, and b) do the numerical values vary? One may hypothesize that if the rank order is roughly similar then there is considerable similarity – and if, in addition, the numerical values are roughly similar, then similarity across countries is established. The reason for starting with rank-ordering is that numerical values might deviate for a number of reasons, while the rank order may be a more robust indicator of the underlying preference structure(s).

Table 9 shows the country-specific rank orderings based on their cardinal values, within the context of all 243 EQ-5D health states, for five randomly selected EQ-5D health states. For example, while health state 11133 was ranked at place number 172 in the Danish set of EQ-5D tariffs, the corresponding health state was ranked at place 65 in the Spanish set of EQ-5D tariffs. It is clear that there are considerable differences - at least for the chosen examples – with the English and Danish set of tariffs being rather close. Focusing on the corresponding numerical TTO cardinal scores, illustrated in Table 9 in square brackets, a considerable disparity is apparent for the health state 23333. The TTO score for Japan was positive while it was negative for the other three countries. In general, though with a few exceptions, Japanese EQ-5D tariffs seem to differ from all the three other countries.

Table 9. Rank ordering for five arbitrarily selected health states. Rank order numbers on a scale 1 to 243. TTO numerical values in square brackets.

Country / *Health state*	**Denmark**	**UK**	**Japan**	**Spain**
11221	10 [0.814]	11 [0.760]	13 [0.729]	9 [0.816]
23221	66 [0.421]	86 [0.208]	87 [0.536]	120 [0.110]
12223	112 [0.270]	107 [0.151]	76 [0.554]	76 [0.247]
11133	172 [0.089]	150 [0.028]	85 [0.541]	65 [0.280]
23333	227 [-0.200]	234 [-0.349]	185 [0.234]	226 [-0.330]

Next the Spearman rank order correlation coefficients were calculated and these are given in Table 10. It is apparent that there is a high degree of correlation, since all coefficients are significant.

Table 10. Spearman correlation matrix between Danish, UK, Japanese, and Spanish EQ-5D tariffs.

	Danish	**UK**	**Japanese**	**Spanish**
Danish	1.000			
UK	0.971*	1.000		
Japanese	0.897*	0.875*	1.000	
Spanish	0.905*	0.920*	0.941*	1.000

Note: *($p < 0.01$).

The Pearson correlation coefficients are presented in Table 11. As with the Spearman coefficients, they (statistically and superficially) indicate a high degree of correlation – but again this probably masks important differences. While Table 9 is only a 'taster', one should probably go through an inspection of all 243 states in a way similar to Table 9. The results are not presented here.

Table 11. Pearson correlation coefficients between Danish, UK, Japanese, and Spanish TTO tariffs.

	Danish	**UK**	**Japanese**	**Spanish**
Danish	1.000			
UK	0.974*	1.000		
Japanese	0.886*	0.856*	1.000	
Spanish	0.919*	0.950*	0.911*	1.000

Note: *correlation significant at ($p < 0.01$).

The results indicate that there existed a certain similarity between European tariffs, and that Japanese tariffs were quite different. While 90 (around 37 per cent) of the Spanish EQ-5D health states were rated as worse than death (i.e. are negative) only 6 (around 3 per cent) of the Japanese EQ-5D health states were rated as worse than death. The equivalent numbers for Denmark and the UK were 54 (around 22 per cent) and 83 (around 34 per cent), respectively.

The results reveal considerable differences in distances of the ranking with 'Denmark' chosen as the *gold standard*. For the UK, 108 (44.4 per cent) of the 243 EQ-5D health states were negative, indicating a lower rank compared to the corresponding ranking for Danish EQ-5D health states. For Japan and Spain, 115 (47.3 per cent) and 114 (46.9 per cent) EQ-5D health states were negative, respectively. Compared to Denmark, 17 of the Japanese EQ-5D health states were ranked more than fifty places higher and 16 were ranked below fifty places. In general, there seemed to be quite a difference in how each country, compared to Denmark, ranked the 243 EQ-5D health states. Illustrations of the above results are presented in Table 12.

Table 12. Distribution of the ranking of the 243 EQ-5D health states in the UK, Japan and Spain, compared to the ranking in Denmark (gold standard). Per cent in square brackets. (n = 243).

Rank differences	UK	Japan	Spain
Higher ranking:			
Above fifty places	0 [0.0]	17 [7.0]	10 [4.1]
Fifty to eleven	53 [21.8]	63 [25.9]	58 [23.9]
Ten to one	61 [25.1]	42 [17.3]	56 [23.0]
Total (positive)	114 [46.9]	122 [50.2]	124 [51.0]
Lower ranking:			
Below fifty places	1 [0.4]	16 [6.6]	10 [4.1]
Fifty to eleven	58 [23.9]	64 [26.3]	68 [28.0]
Ten to one	49 [20.2]	35 [14.0]	36 [14.8]
Total (negative)	108 [44.4]	115 [47.3]	114 [46.9]
Match (same ranking)	21 [8.6]	6 [2.5]	5 [2.1]

VAS versus TTO EQ-5D valuations

All 1,332 individuals valued the same set of EQ-5D health states using both the VAS and the TTO techniques. This offered a unique opportunity to investigate possible differences between the two methods assessed by the same (sample) population. The same regression techniques were applied on the VAS data as used to estimate EQ-5D TTO based tariffs (cf. figure 4). Surprisingly, this resulted in a different choice of best-fitting model than in the case of estimating non-direct TTO valuations. The best-fitting model at predicting non-direct TTO valuations (i.e. TTO3) resulted in both inconsistent and insignificant parameters when used to predict non-direct VAS valuations. So did the TTO4 model. Out of the eleven models tested only two models, TTO1 and TTO2, resulted in both consistent and significant parameters (at the 1 per cent significance level). All eleven models yielded a high R^2 value. Unfortunately, this meant that the best-fitting model in predicting non-direct VAS valuations differed from the best-fitting model used to predict non-direct TTO valuations. The implication was that differences in the two sets of tariffs might be due, to some degree, to selection bias (i.e. model specification). To what degree this was the case is impossible to measure. In making the selection bias as low as possible the TTO1 model, which excluded the N3 term (cf. Figure 4) was applied, since this model specification was closer to the TTO3 specification than was the TTO2 model. R^2 was higher in TTO2 compared to TTO1, however, only moderately. The estimated parameters used to estimate tariffs based on the non-direct VAS valuations are illustrated in Table 13. The model specification is referred to as VAS1 and resembles the TTO1 model specification.

Table 13. Estimates of model VAS1 used in the valuation exercise.

Estimates	VAS1
Intercept	0.1362 (<0.0001)
MO	0.0566 (<0.0001)
SC	0.0690 (<0.0001)
UA	0.0870 (<0.0001)
PD	0.1398 (<0.0001)
AD	0.1678 (<0.0001)
R^2	0.7462

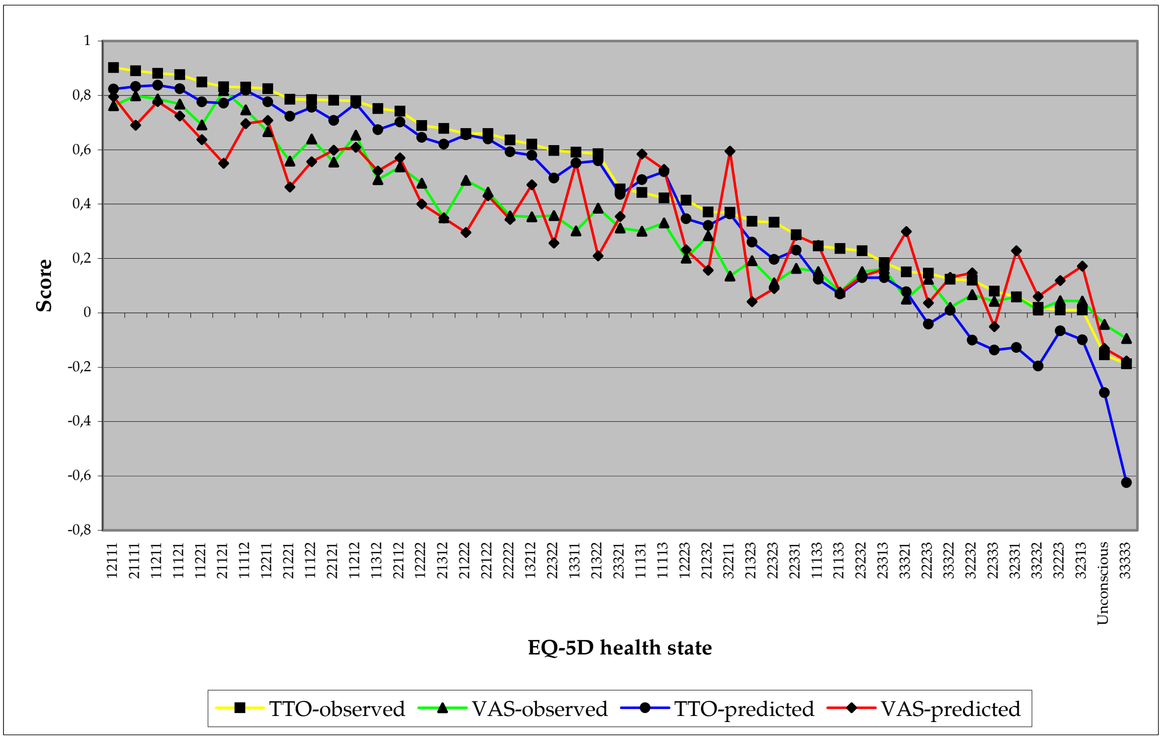
A set of VAS based EQ-5D tariffs based on the estimates in Table 13 was estimated and compared to the EQ-5D tariffs modelled using the TTO3 model. The results are shown in appendix D. The numeric correlation and rank-ordered correlation of the two sets of tariffs were assessed using both Pearson and Spearman correlation coefficients. Not surprisingly, both correlation coefficients came out fairly high and significant at the 1 per cent level; Pearson = 0.855 and Spearman = 0.847. Looking at the differences between the two sets of tariffs, 48 (20 per cent) of the 243 health states in the TTO-based EQ-5D tariff were negative, as were 20 (8 per cent) of the VAS EQ-5D health states. For 125 health states (around 50 per cent) the TTO-based EQ-5D tariffs were higher than the VAS-based EQ-5D tariffs, with the majority being severe health states (judged by at least one '3').

Comparison between direct and predicted valuations

Valuations of the 46 health states, which the respondents valued directly – both in the VAS and the TTO exercise – as well as the predicted VAS and TTO valuations, were compared. The mean observed valuations of the 46 directly valued states using both the VAS and TTO methods are illustrated in figure 6, alongside with the predicted valuations.[5]

[5] Predicted valuations were based on model TTO3 for the TTO exercise and VAS1 for the VAS exercise.

Figure 6. Mean TTO and VAS values for 46 EQ-5D health states – observed (rescaled) and predicted. (n = 1,332).



Note: Predicted valuations were based on model TTO3 for the TTO exercise and VAS1 for the VAS exercise.

As seen in the above figure the observed TTO valuations were higher than the corresponding VAS valuations in the case of mild and moderate health states, but lower for the more severe health states, i.e. health states worse than death. VAS values covered a considerably smaller range of the potential valuation space than the TTO valuations. In general, the same scenario could be seen with regard to the predicted valuations. However, where the predicted TTO valuations decreased regularly as the health states became more severe, there were several unexplained movements in the predicted VAS valuations.

Estimating a power function

As noted by Nord (1991), it is unclear what people mean when they value health states on a VAS. Comparisons of VAS valuations with valuations elicited by means of the equivalence of numbers technique, indicate that intervals between states on the VAS must be weighted more the closer they are to the bottom of the scale. This is the same as saying that VAS values should not be used directly as utility weights for life years. A transformation of the values is needed, and a *power function* may be suitable for this purpose.

Even though observations were at the individual level, it was impossible to estimate a power function, due to the many TTO based EQ-5D health state valuations worse than dead, i.e. negative valuations. Hence data at the aggregated level, i.e. mean values for all 46 EQ-5D health states were applied instead. Bleichrodt & Johannesson (1997) investigated empirically whether a stable relationship between VAS and SG valuations existed. They concluded that *"The results of the estimation procedure using individual data were not supportive of a stable relationship: in general, the parameters obtained differed significantly, we found indications of misspecification, and in most cases there was indication of heteroskedastic errors."* Torrance *et al.* (2001) commented on the findings reported by Bleichtrodt & Johannesson: *"This is not surprising, since it has long been known that the relationship may not be stable at the individual level."*

It has long been known that VAS valuations do not agree with either TTO or SG valuations, respectively. This was first reported in Torrance (1976), where he noted that VAS valuations (mean scores for health states) were consistently lower than TTO valuations. Such findings have since then been reported by others [Read *et al.* 1984; Elstein *et al.* 1986; Boyd *et al.* 1990; Bass *et al.* 1994]. In general, the assumption is that there is the following relationship between VAS, TTO, and SG values: VAS<TTO<SG [Krabbe *et al.* 1997].[6]

The re-scaled EQ-5D health states from the VAS exercise, where the range was from zero (worst) to 1 (best), were applied. The data from the TTO exercise were in the range of -1 (worst) to 1 (best). In *SAS*, the following expression: log (1 – VAS) = a * log(1 – TTO), which is equivalent to the expression suggested by Torrance *et al.* (1976): VAS = 1 – $(1 – \text{TTO})^a$ was regressed.[7] It was found, based on 46 means, where the number of respondents was 1,332, that a = 0.66 (adjusted R^2 = 0.87), so resulting the power function to take on the following expression: VAS = 1 – $(1 – \text{TTO})^{0.66}$.[8]

Several studies have tried to estimate a similar power function between VAS and TTO EQ-5D health state valuations in order to transform VAS values into TTO utilities. Torrance *et al.* (1997) reported a power function with a coefficient of 0.62 (R^2 = 0.80) based on 18 means of valued health-state scenarios (number of respondents app. 200). A study by Stiggelbout *et al.* (1996) reported a coefficient of 0.64 and Loomes (1993) reported a coefficient of 0.55 based on a secondary analysis of data by Bombardier *et al.* (1982). Krabbe *et al.* (1997) found the coefficient, based on 13 mean values, to be 0.42 (R^2 = 0.96) Finally Busschbach reported, in Krabbe *et al.* (1997), similar results, namely a coefficient of

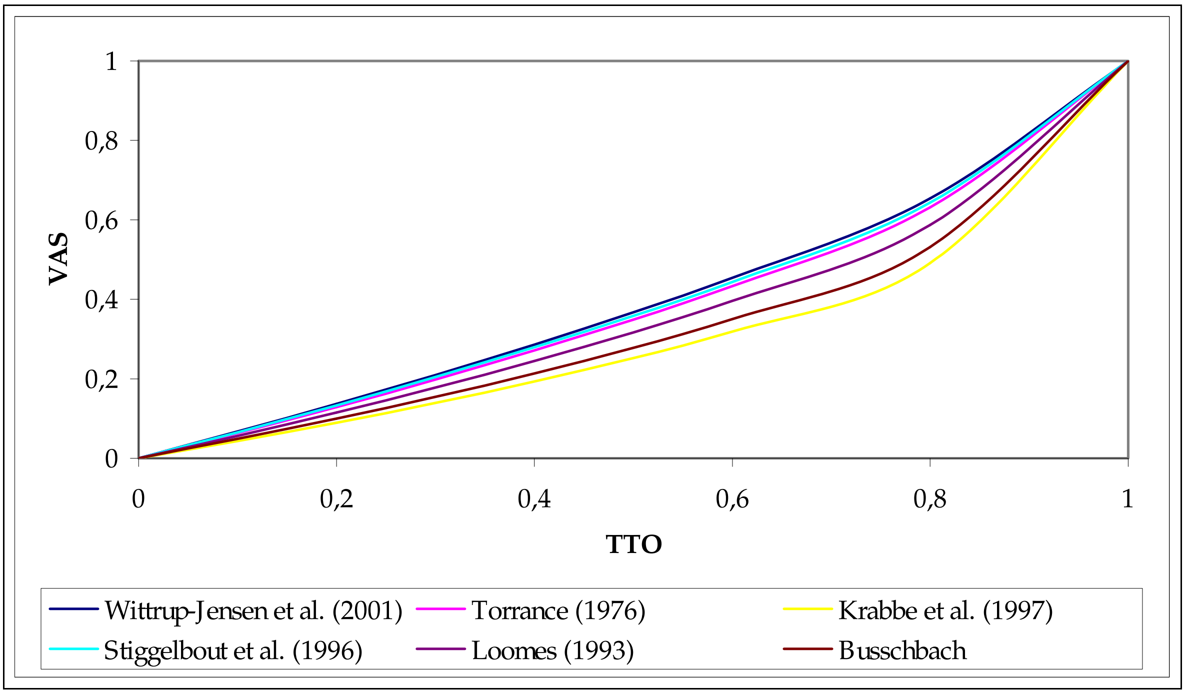
[6] While this is the state-of-the-art assumption, some studies have shown, concerning the relationship between VAS and the TTO, that VAS values are > TTO values for mild to moderate health states and < for severe health states [Badia *et al.* 1998; Dolan & Sutton 1997].

[7] Actually, Torrance's (1976) original relationship between VAS and SG valuations was u = $1-(1-v)^a$, which expresses the concave relationship between VAS and SG valuations. In this study the aim was to estimate VAS valuations as the dependent variable and TTO valuations as the independent variable, which results in a convex relationship.

[8] The regression technique we apply is based on a regression without an intercept. If we include the intercept, a is 0.66 (adjusted R^2 = 0.74). From a purely statistical point of view the model should be estimated without an intercept, but it is not totally clear whether this is enough to exclude the intercept.

0.47 (R^2 = 0.95) for 103 respondents. The estimated power functions between VAS and TTO valuations documented within the literature and the findings are reported and illustrated in figure 7.

Figure 7. Relationship between mean TTO and VAS valuations for EQ-5D health states.



Discussion

I. Estimating a Danish set of EuroQol (EQ-5D) tariffs

In this study we employed econometric analysis based on a regression analysis where the dependent variable was (one minus) the score given to EQ-5D health states. All independent variables were expressed as dummy variables and as such derived from the ordinal nature of the EQ-5D descriptive classification system. The functional form of the regression was linear additive, which seems reasonable given the assumption that valuations within the TTO method exhibit interval scale properties. According to Mehrez & Gafni (1990) there is no formal theoretical foundation for the TTO technique (as there is for the SG) and it is highly questionable whether individuals participating in TTO studies have value functions consistent with these assumptions. While there exist empirical studies that support Mehrez & Gafni (Cook *et al.* 2001), the TTO technique is still applied and its results interpreted as utilities (and not merely just values) in many studies. More empirical work is required on whether the TTO method expresses utility, and hence TTO valuations can be interpreted as utilities on an interval scale, i.e. a method that elicits individuals' preferences for health.

We did not apply or test functional forms other than the additive model since this, according to Dolan *et al.* (1995), would be difficult given the nature of the independent variables. To our knowledge no similar studies have tried to apply other kinds of functional forms or to come up with any theoretical justification for further transformation of the valuations themselves. Since existing studies of national EQ-5D tariffs have more or less applied the regression specification based on *ad hoc* assumptions, we formulated different models and tested these against each other, specifically looking at models that yielded consistent and significant parameters.

Our study was conducted at the individual level, where each individual valued a discrete number of EQ-5D health states. Since individuals valued more than one, and also different, health states the error term in the regression analysis was partly explained by the individual herself/himself. This violated one of the assumptions for applying an Ordinary Least-Square (OLS) regression technique. Hence, we employed a random effect (RE) model technique, which explicitly has an individual specific error term representing the extent to which the intercept for the *i*'th individual differs from the overall intercept α. We also specified a special case of the RE model called the random coefficient (RC) model, in which only the intercept is randomised, to see whether this model was superior to the RE model. The RC yielded inconsistent parameters as some were negative. The split-sample test showed a strong significant difference between the RC and the RE parameters, which indicated efficiency problems in the specified RE model. However, since the RC parameters were inconsistent we chose to go with the RE model.

Even though we rejected the OLS model on purely theoretical grounds we also tested the performance of the OLS model empirically. The OLS was tested against the RE model using three tests: the Likelihood ratio (LR) test, the Wald test and the Lagrange Multiplier (LM) test. All three tests rejected the OLS model compared to the RE model. For the present study an implicit implication of using the RE model is that individual deviations form the average valuations of health states did not vary over the range of health states. In other words, if an individual overestimates each of the health states presented by a certain amount, then this amount should not change if the individual is presented with any other health state. This assumption cannot explicitly be given, based on pure *ad hoc* assumptions, but needed to be tested. A fixed effects (FE) model was applied where a violation of this independence assumption would result in consistent parameters. We tested the FE model against both the RE model and the OLS model using a Hausman test. The Hausman test rejected both the RE and the OLS models. Furthermore, the FE model performed better concerning the significance of the parameters, but at the same time yielded inconsistent parameters. Hence we chose the RE model.

The RE model failed the Ramsey RESET test indicating that the RE model suffered from misspecification and a different test showed that the RE model also suffered from heteroscedasticity. The RE model did not, however, suffer from multicollinearity, since the conditional number (CN) was 12.21 (the critical value, according to Green (2000), is around 20-30). The fact that the models suffered from misspecification is not surprising since, as reported by Dolan (1997), the power of the RESET test increases as the sample size increases. Thus, with a total of over 18,000 observations (the number of individuals multiplied by the number of health states valued directly), any model with relatively few independent variables is likely to be mis-specified. The problems associated with the presence of heteroscedasticity are also difficult to overcome, since transformation of one or more independent variable is not feasible, given the (categorical) nature of these variables. As noted by Dolan (1997) heteroscedasticity is likely to result in inefficient rather than biased parameters and so, even though misspecification is a problem, there is little which can be done. The same problems concerning misspecifications and heteroscedasticity are reported in similar studies [Dolan 1997; Busschbach *et al.* 1999; Devlin *et al.* 2000]. However, as reported in Busschbach *et al.* (1997), the misspecification disappears when health states are presented randomly.

The direct valuations of the EQ-5D health states were strongly skewed to the right i.e. many health states were valued around +1. In order to compensate for this relationship, we applied a *censored Tobit model*. However, surprisingly, this model resulted in inconsistent parameters. The problem with skewed data still remains and we would like to believe that the censored Tobit model is the most accurate for predicting non-direct EQ-5D valuations. A possible explanation for the censored Tobit model resulting in inconsistent parameters is that the model somehow ignores the censoring of the 'good' health states, i.e. EQ-5D health states with an over weighting of 1's and 2's, and puts too much emphasis on the decline from 2 to 3.

To summarize, we chose the RE model over the FE model, thus implicitly rejecting the existence of socio-economic characteristics with a significant influence on how individuals value health (states). However, it has to be stressed that the FE model is, at least from a purely statistical point of view, superior to the RE model, although the FE model resulted in inconsistent parameters. Rejecting the FE model implies that individual characteristics did not vary across the valuation of health (states). However, explicitly rejecting that socio-economic characteristics may be irrelevant in how individuals value health (states), is a strong postulation and it needs to be addressed more deeply.

Compared to the MVH study in the UK, we have been much more explicit in our modelling. We do not intuitively rely on the RE model to be the most appropriate regression model. Although we ended up using the RE model, our extensive testing showed that it is important to be very explicit about which model to use. We urge that future studies that estimate national EQ-5D tariffs blue print our testing, since our findings may not be replicating if different data are applied.

Qualitative studies connected with the MVH study (Kind *et al.* 1994) suggest that some of the observed differences in valuations might result from older respondents doubting the plausibility of the 'worse than death' scenario. Dolan (2000) investigated whether *age* is a significant factor in determining valuations for none-directly valued EQ-5D health states. The results showed that a tariff based on values from those aged 60 and over was considerably lower than a tariff based on values from those aged 18-59 years. After adjusting the scores of older individuals, based on the assumption that their lower values for health states rated as worse than death is due to an experimental artefact, the differences between the valuation tariffs of the two age groups was greatly diminished. Dolan (2000) concluded that age may have an important effect on health state valuation. Although our study was randomised across age groups, we experienced an over-representation of the age group ≥ 60 years compared to the general population. As we decided not to weight the sample data it may be the case that this group's valuations influence the regression estimates so that the estimates are underestimated. However, this is only speculation and it needs to be tested more formally in an empirical study. One way is simply to replicate the studies presented in Dolan (2000). Furthermore, there is a need for studies that investigate whether and how other socio-economic characteristics such as gender, education, and social status influence the valuation of EQ-5D health states.

Using the literature we applied several models which past studies also have tested. Nevertheless we ended up with the 'best-fitting' model being different from that which past studies report. National studies conducted in England [Dolan 1997], Japan [Fukuda et al. 1999], and Spain [Badia et al. 2001] all report the same model, similar to our model TTO4, as the one to be used in estimating valuations for non-directly assessed EQ-5D health states, but this is not necessarily the best-fitting model. Fukuda *et al.* (1999) reported estimates based on a model, similar to our TTO4 model, where the *p*-value for the N3 variable was insignificant ($p = 0.284$). Furthermore, they reported that if they omitted the N3 variable, all parameters were consistent and significant at the 1% level. We believe, even though this is not reported anywhere in the literature, that the model originally reported by the MVH Study Group is not necessarily is the best-fitting model for the English data, and that the N3 term is added based on purely intuitive grounds. In addition, we suspect that both the Spanish and Japanese studies tried to fit their data into a model including the N3 term following the intuitive postulations put forward in the English study. In Dolan (1997) the term N3 is defined as an intercept dummy where any of the dimensions were at level 3. Dolan continues *"Without this additional dummy, which can be interpreted as reflecting the much greater disutility associated with "extreme problems", the residuals were related systematically to the predicted values in that the model underestimated the values of less severe states and overestimated the values of more severe ones."* This supports our contention that the MVH Study Group included the N3 term based on theoretical speculations and not based on which model was the best-fitting model judged by consistent and significant parameters. We support Dolan (1997) in that there may exist an intuitive explanation for including the N3 term, however our results show that in-

cluding the term N3 makes one of the other parameters insignificant, and hence we chose the TTO3 model to estimate the Danish EQ-5D tariffs based on this model specification.
Devlin *et al.* (2000) specified an N2 term, which should account for any movement from level 1 to 2. We tried to incorporate the N2 term in the models TTO10 and TTO11. While TTO10 resulted in the N2 term being insignificant at the 10% level, in TTO11 the N2 term was inconsistent (i.e. negative). The reason for both models failing is likely to be 'double-counting', since the N2 term and the intercept cover the same thing. The intercept covers any movement away from perfect health, where any movement away from full health begins with (at least) one level across the dimensions being 2. If one wishes to incorporate the N2 term one should experiment with leaving out the intercept or simply leaving out the N2 term from the beginning. Finally, it appears that all models (TTO1 to TTO11) fitted the data rather well in terms of R^2. However, as noted by Busschbach *et al.* (1999), the R^2 term is difficult to interpret and we believe that all models are more or less subject to some kind of misspecification.

As noted by Dolan & Roberts (2001), the EQ-5D tariffs are used to express the value of differences between health states and thus it would be worth exploring whether a tariff could be calculated using differences between values rather than using the values themselves.[9] These authors used the data from the MVH Study and estimated EQ-5D tariffs based upon the differences in value between the worst possible state (i.e. 33333) and all other states. They concluded that the new model (the new set of estimated tariffs) had a better predictive ability than the original and that it produced some differences that might be considered important in evaluative studies. Hence they recommended that the new model should be used alongside the original to see what practical differences it might make. It is difficult to say anything conclusive concerning whether tariffs ought to be based on absolute values or differences between values. However, the authors' point is worth pursuing. The next step could be to use our dataset to estimate a set of EQ-5D tariffs based on the differences between values rather than using the values themselves. However, we do not agree with what they recommend, i.e. that both sets should be used, as there is a danger that one would simply apply the set which best suits the study in question. We believe that one of the sets should be used but not both, and that the choice should be based on *a priori* reasoning.

According to Dolan (1997b) little attention has been paid to how individual responses should be aggregated when attempting to express the valuations of a given group. The modelling in our study is based on individual data and the EQ-5D tariffs are (as in other documented TTO studies) an approximation of *mean* values. Would the tariffs differ if the tariffs were to be based on *median values* and what are the implications? Dolan (1997b) used the data from the MVH study to elicit English tariffs based on median values and then compared them to the original tariffs based on mean values. The results were that the median-based tariffs, compared to the mean-based tariffs, have higher values for

[9] Permission for citation obtained from the authors.

less severe EQ-5D health states and lower values for more severe health states. Dolan (1997b) concluded that this is likely to have important implications for resource allocation decision of scarce health care services. Since neither theoretical indicators nor guidelines exist to decide whether to use individual-based or median-based tariffs, the tariffs to be used should be based on a matter of judgment. However, it has to be stressed that the choice ultimately should be based upon a prior theoretical position rather than on intuition about which set of tariffs seems to produce better answers.

II. Comparison of Danish EQ-5D tariffs with other national EQ-5D tariffs

One of our aims in this study was to compare our TTO tariffs with similar studies across other countries in order to see similarities or differences that can speak for or against a cross-country TTO tariff. Using a similar regression model we estimated a set of tariffs and compared these with tariffs estimated in the UK, Japan and Spain. Not surprisingly both the Pearson and Spearman correlation matrices showed high and significant correlations, which indicates that tariffs are similar across different countries and that it does not matter whether, for example, English or Danish tariffs are applied in Danish economic evaluation (e.g. cost-utility analysis). However, a more direct comparison of the tariffs using a rank ordering of the health states and their numerical values shows that some differences exist. While around 37 per cent of the Spanish health states are rated as worse than death, the figure is only 3 per cent for Japanese health states. The same figures for Denmark and the UK are 22 and 34 per cent, respectively. The implication is that caution should be taken in concluding that it does not matter which national tariffs are applied. In order to say something conclusive, a cost-utility anakysis should be conducted, applying all four sets of tariffs to see if there are any similarities/differences. The discussion to what extent it matters whether Danish or English tariffs are used should not only take place in an empirical context, but also on an intuitive level. Even if the differences were to be minor between Danish and UK tariffs, the intuitive argument would be that tariffs ought to be country-specific. The direct implication is that every country ought to have a set of EQ-5D tariffs estimated within its own national population. This is a very strong postulation that needs more research.

A possible next step could be to pool individual TTO valuations across the four countries in order to investigate similarities/differences. Badia et al. (2001) reported on a pooling between the English and Spanish sets of TTO data. For the milder health states Spanish and UK values were similar. For intermediate health states, Spanish values were both higher and lower than UK values, whereas for health states worse than death, English values were in general higher than Spanish values. Furthermore, they reported that there were statistically significant differences in 34.9 percent of the health states valued directly. It seems that UK raters ascribed greater importance to the dimensions of 'pain/discomfort' and 'anxiety/depression', whereas Spanish raters placed more importance on the functional dimensions of 'mobility' and 'self-care'.

One-way to illustrate whether it matters to apply EQ-5D tariffs from other countries in a Danish context is to perform a cost-utility analysis. Even though the correlation coefficients were significant at the 1% level, applying country-specific tariffs in the CUA example resulted in quite different cost-utility ratios. However, this is a stand-alone example and these findings may not be reproduced in other cost-utility studies in other disease areas. Nevertheless, the conclusion is that from an empirical perspective it does matter whether Danish, English, Japanese, or Spanish EQ-5D tariffs are used as weights and hence this may have an impact on the allocation of health care resources.

III. Comparison of TTO- and VAS-based EQ-5D valuations

Since all respondents valued the EQ-5D health states using both the VAS and the TTO technique, this offers a unique situation for a direct assessment of the valuation techniques. Using the 46 health states which are directly valued at the individual level, aggregated data (group means), and finally the estimated sets of tariffs, we looked for any patterns between the two valuations. Torrance (1976) was the first to attempt to construct conversion curves to relate different scaling methods. In the case of the VAS method versus the TTO method he concluded *"… the two techniques exhibit a systematic relationship [which] can be approximated by a number of different functions. Two that fit well…are a logarithmic function and a power function"*. This conclusion was based on aggregate data but did not hold for the individual level. Other researchers obtained similar, but yet different, results. Wolfson *et al.* (1982), based on an estimated linear relationship, concluded that SG values were much higher than VAS valuations, with TTO values generally somewhere in between. Results obtained from Read *et al.* (1984) support these findings. Hornberger *et al.* (1992) found that TTO produced the highest mean values followed by VAS and then SG. The latter findings contradicted the standard VAS<TTO<SG relationship as reported, for example, by Krabbe *et al.* (1997). However, it should be noted that the study conducted by Hornberger *et al.* (1992) invoked patients' valuations of their own health, whereas other studies are based on hypothetical scenarios. Empirical studies reported by Badia *et al.* (1999b) show that by using median valuations, TTO valuations were higher than VAS valuations for mild health states and lower for more severe health states. These findings are supported by Dolan & Sutton (1997), who also found that TTO valuations (including the use of props) were higher than VAS valuations for mild health states and lower for more severe health states. However, they also reported that 'no props' responses (i.e. TTO responses) were consistently higher than VAS valuations, particularly for more severe health states. Using the observed aggregated (group means) health state valuations we found that TTO valuations were higher than the corresponding VAS valuations in the case of mild and moderate health states, but lower for more severe health states, i.e. health states worse than death. The scenarios of the predicted valuations are roughly the same. Thus our findings support the results reported by both Badia et al. (199b) and Dolan & Sutton (1997).

Surprisingly, the model TTO3 results in both inconsistent and insignificant parameters when we apply it on the valuations from the valuation exercise. Only model TTO1 and TTO2 come up with consistent and significant parameters. In order to reduce bias, we chose TTO1, that is, excluding the N3 term to estimate the predicted VAS valuations. Having estimated a set of tariffs based on both the VAS and TTO exercises, we compared the two sets. Both the Pearson and Spearman correlation coefficients were significant at the 0.01 level, indicating a high correlation between the two sets of tariffs. However, looking more closely at direct comparison of VAS and TTO tariffs revealed huge differences in the scores. While around 20 per cent of the TTO scores were negative (worse than death), around 8 per cent of the VAS scores were negative. Around 50 per cent of the TTO states had a higher score than the corresponding VAS states, where the majority were *severe* health states.

IV. Estimating a power function

As suggested by Torrance (1976), Torrance et al. (2001) and Nord (1991), VAS valuations need to be transformed to TTO utilities in order to be usable. First we would like to stress that we seriously doubt that such an exercise transforms valuations into utilities (based on the grounds that TTO valuations do express utilities). A power function merely expresses a linear fit between the two valuation methods across health states and therefore, based on theoretical grounds, cannot be used to transform *values into utilities*.[10] Nevertheless, we used the rescaled VAS and TTO valuations of the 46 directly valued health states at the aggregate level (mean values) to estimate a power function based on the structure $VAS = 1 - (1 - TTO)^a$, suggested by Torrance (1976), where we have put the VAS valuations as the dependent variable and the TTO valuations as the independent variable. Given the structure of the power function, we initially did a log transformation of the valuations and then modelled the valuations within a linear regression analysis, without including an intercept, since this is the intuitively correct thing to do. This resulted in $a = 0.66$ (adjusted $R^2 = 0.87$). Our findings are fairly close to similar power functions estimated between VAS and TTO valuations. However, the causes of these differences can be manifold.

As the form of the power function, which is applied here to illustrate the functional relationship between VAS and TTO tariffs, is based on a more or less *ad hoc* assumption originally suggested by Torrance (1976) concerning how the relationship between the two valuation methods could be described, this does not necessarily mean that this approach is the right one from a normative perspective. Instead of applying a power function to transform the VAS valuations, we investigated whether a regression analysis (e.g. a *Box-Cox transformation*) would be more applicable. However, given the nature

---

[10] This issue is discussed in more detail in Chapter 3, where we apply the power function to VAS valuations.

and form of the data, that is, many negative valuations, we were unable to locate a function that could be fitted to the data, without either rescaling or excluding negative valuations.

In conclusion, this chapter presents a national randomized and representative EQ-5D set of tariffs based on the TTO scaling technique, which can be used as an input for Danish cost-utility analyses where the EQ-5D profile is applied. Although we also estimate corresponding VAS tariffs, we would like to stress that these should not be applied in Danish cost-utility analyses, due to their lack of foundation in axiomatic decision theory.[11] The estimation of VAS tariffs in this chapter merely illustrates the empirical relationship that exists between the TTO and VAS techniques.

11 See also Bleichtrodt & Johannesson (1997).

References

1. Austin PC, Escobar M, Kopec JA. The use of the Tobit model for analyzing measures of health status. *Quality of Life Research* 2000; 9: 901-910.
2. Badia X, Herdman M, Kind P. The influences of ill-health experience on the valuation of health. *Pharmaeconomics* 1998; 6: 687-96.
3. Badia X, Monserrat S, Roset M, Herdman M. Feasibility, validity and test-retest reliability of scaling methods for health states: The visual analogue scale and the time trade-off. *Quality of Life Research* 1999a; 8: 303-310.
4. Badia X, Roset M, Herman M. Inconsistent responses in three preference-elicitation methods for health states. *Social Science and Medicine* 1999b; 49: 943-950.
5. Badia X, Roset M, Herdman M, Kind P. A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. *Medical Decision Making* 2001; 21: 7-16.
6. Bennet K, Torrance GW, Tugwell P. Methodological challenges in the development of utility measures of health-related quality of life in rheumatoid arthritis? *Controlled Clinical Trials* 1991; 12.
7. Berzon R & Shumaker S. A critical review of cross national health-related quality of life instruments. *Quality of Life Newsletter* 1993; 5: 1-2.
8. Bombardier C, Wolfson AD, Sinclair AJ, McGreer A. comparison of three preference measurement methodologies in the evaluation of a functional index. In Deber R & Thompson G (eds.) *Choices in health care: Decision making and evaluation of effectiveness*. University of Toronto: Toronto. 1982.
9. Brooks R, with the EuroQol Group. EuroQol: The current state of play. *Health Policy* 1996; 37: 53-72.
10. Buckingham K & Drummond M. *A theoretical and empirical classification of health valuation techniques*. HESG Conference: Strathclyde. 1993.
11. Bullinger M. Quality of life: Definition, conceptualisation and implications – a methodologist's view. *Theoretical Surgery* 1991; 6: 143-148.
12. Bush JW, Kaplan RM, Berry CC. Comparison of methods for measuring social preferences for a health state index. American Statistical Association Annual Meeting 1997: 682 -.
13. Busschbach JJV, McDonnell J, van Hout BA. Testing different parametric relations between the EuroQol health description and health valuation in students. In Nord E. (eds.) *Conference*

proceedings of the EuroQol plenary meeting. Working Paper No. 2/97. National Institute of Public Health: Oslo. 1997.

14. Busschbach JJV, McDonnell J, Essink-Bot ML, van Hout BA. Estimating parametric relationships between health description and health valuation with an application to the EuroQol EQ-5D. *Journal of Health Economics* 1999; 18: 551-571.
15. Carr-Hill RA. Allocating resources to health care: Is the QALY (Quality Adjusted Life Years) a technical solution to a political problem? *International Journal of Health Services* 1991; 21: 351-363.
16. Cook KF, Ashton CM, Byrne MM, Brody B, Geraci J, Giesler BR, Hanita M, Souchek J, Wray NP. A psychometric analysis of the measurement level of the rating scale, time trade-off, and standard gamble. *Social Science and Medicine* 2001; 53: 1275-1285.
17. Dolan P, Gudex C, Kind P, Williams A. *A social tariff for EuroQol: Results from a UK general population study*. Discussion Paper 138. Centre for Health Economics. 1995. The University of York.
18. Dolan P. The effect of experience of illness on health state valuations. *Journal of Clinical Epidemiology* 1996; 49: 551-64.
19. Dolan P, Gudex C, Kind P, Williams A. Valuing health states: A comparison of methods. *Journal of Health Economics* 1996a; 15: 209-231.
20. Dolan P, Gudex C, Kind P, Williams A. The time trade-off method: Results from a general population study. *Health Economics* 1996b; 5: 141-154.
21. Dolan P & Kind P. Inconsistency and health state valuations. *Social Science and Medicine* 1996; 4: 609-615.
22. Dolan P. Modelling valuations for EuroQol health states. *Medical Care* 1997a; 35(11): 1095-1108.
23. Dolan P. Aggregating health state valuations. *Journal of Health Services Research and Policy* 1997b; 2(3): 160-165.
24. Dolan P. Whose preferences count? *Medical Decision Making* 1999; 19: 482-486.
25. Dolan P. The measurement of health-related quality of life. In Culyer AJ & Newhouse JP (eds.) *Handbook of health economics*. Elsevier North-Holland: Amsterdam. 2000a.
26. Dolan P. Effect of age on health state valuations. *Journal of Health Services Research* 2000; 5(1): 17-21.
27. Dolan P & Roberts J. *Modelling valuations for EQ-5D health states: An alternative model using differences ini valuations*. Paper presented at iHEA Conference, York, July 2001.
28. Festinger L. *A theory of cognitive dissonance*. Stanford University Press. 1957.
29. Froberg DG & Kane RL. Methodology for measuring health-state preferences – II: Scaling methods. *Journal of Clinical Epidemiology* 1989a; 42: 459-471.

30. Froberg DG & Kane RL. Methodology for measuring health state preferences – III: Population and context effects. *Journal of Clinical Epidemiology* 1989b; 42: 585.
31. Froberg DG & Kane RL. Methodology for measuring health-state preferences – I: Measurement strategies. *Journal of Clinical Epidemiology* 1989c; 42: 345-354.
32. Fukada T, Hamashima C, Hisashige A, Ikeda S, Ikegami N, Nishimura S, Sakai I, Tamura M, Tsuchiya A. *Estimating a EuroQol value set: The case of Japan*. Discussion Paper presented at the EuroQol Group meeting in 1999, Sitges, Spain.
33. Gafni A, Birch S, Mehrez A. Economics, health and health outcomes: HYEs versus QALYs. Journal of *Health Economics* 1993; 11: 325-339.
34. Gescheider GA. Psychological scaling. *Annual Review of Psychology* 1988; 39: 169-200.
35. Giesler RB, Ashton CM, Brody B, Byrne MM, Cook K, Geraci JM, Hanita M, Souchek J, Wray NP. Assessing the performance of utility techniques in the absence of a gold standard. *Medical Care* 1999; 37: 580-588.
36. Gold MR, Patrick DL, Torrance GW, Fryback DG, Hadorn DC, Kamlet MS, Daniels N, Weinstein M. Identifying and valuing outcomes. In Gold MR, Siegel JE, Russell LB, Weinstein MC (eds.) *Cost-effectiveness in health and medicine*. Oxford University Press: New York. 1996.
37. Greene WH. *Econometric analysis*. Fourth Edition. Printice Hall International: New York. 2000.
38. Green C. On the societal value of health care: What do we know about the person trade-off technique? *Health Economics* 2001; 10: 233-243.
39. Gudex C, Dolan P, Kind P, Williams A. Health state valuations from the general public using the visual analogue scale. Quality of Life Research 1997; 5: 521-31.
40. Hadorn DC. The role of public values in setting health care priorities. *Social Science and Medicine* 1991; 32: 773-781.
41. Hornberger JC, Redelmeier DA, Petersen J. Variability among methods to assess patients well-being and consequent effect on a cost-effectiveness analysis. *Journal of Clinical Epidemiology* 1992; 45(5): 505-512.
42. Johannesson M, Jönsson B, Karlsson G. Outcome measurement in economic evaluation. Health Economics 1996; 5: 279-296.
43. Kahnemann D & Tversky A. Prospect theory: An analysis of decision under risk. *Econometrica* 1979; 4: 580-588.
44. Kahnemann D & Tversky A. The framing of decisions and the psychology of choice. *Science* 1981; 211: 453.
45. Kind P, Gudex C, Dolan P, Williams A. Practical and methodological issues in the development of the EuroQol: The York experience. *Advances in Medical Sociology* 1994; 5: 219-253.
46. Krabbe PFM, Essink-Bot ML, Bonsel GJ. The comparability and reliability of five health-state valuation methods. *Social Science and Medicine* 1997; 45(11): 1641-1652.

47. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, Ciampi A, Till JE, Boyd NF. The measurement of patients' values in medicine. *Medical Care* 1982; 17: 501.
48. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, Ciampi A, Till JE, Boyd NF. Describing health states: Methodologic issues in obtaining values for health states. *Medical Care* 1984; 22: 543.
49. Loomes GL & Sugden R. Regret theory: An alternative theory of rational choice under uncertainty. *Economic Journal* 1982; 92.
50. Loomes GL & McKenzie L. The use of QALYs in health care decision making. *Social Science and Medicine* 1989; 28: 299-308.
51. Loomes G. Disparities between health state measures: Is there a rational explanation? In Gerrard B (eds.) *The economics of rationality*. Routledge: London. 1993.
52. Mehrez A & Gafni A. Quality-adjusted life years, utility theory, and health-years equivalents. *Medical Decision Making* 1991; 11(2).
53. Merbitz C, Morris J, Grip JC. Ordinal scales and foundations of misinference. *Archives of Physical and Medical Rehabilitation* 1989; 70: 308-312.
54. Mulley AG. Assessing patient's utilities. Can the ends justify the means? *Medical Care* 1989; 27: 269-281.
55. Nord E. Methods for quality adjustment of life years. *Social Science and Medicine* 1992; 34: 559-69.
56. Ohinmaa A & Sintonen H. *Inconsistencies and modelling of the Finnish EuroQol (EQ-5D) preference values*. Paper presented at the 15th yearly EuroQol Group Meeting in Hannover, 1 – 2 October 1998.
57. Patrick DL, Starks HE, Cain KC, Uhlmann RF, Pearlman RA. Measuring preferences for health states worse than death. *Medical Decision Making* 1994; 14: 9 – 18.
58. Pliskin JS, Shepard DS, Weinstein MC. Utility functions for life years and health status. *Operations Research* 1979.
59. Prades JLP. Is the person trade-off a valid method for allocating health care resources? *Health Economics* 1997; 6: 71-81.
60. Rabe-Hesketh S & Everitt B. *A handbook of statistical analyses using STATA*. New York: CRC Press. 2000.
61. Read JL; Quinn RJ; Berrick DM; HV Fineberg, Weinstein ML. Preferences for health outcomes: Comparison of assessment methods. *Medical Decision Making* 1984; 4(3): 315-329.
62. Richardson J. Cost utility analysis: What should be measured? *Social Science and Medicine* 1994; 39: 7-21.
63. Roest FH, Eijkemans MJ, van der DJ, Levendag PC, Meeuwis CA, Schmitz PI, Habbema JD. The use of conficence intervals for individual utilities: Limits to formal decision analysis for treatment choice. *Medical Decision Making* 1997; 17: 285-291.

64. Rutten-Mölken MP, Doorslaer EKA, Vliet RCJA. Statistical analysis of cost outcomes in a randomized controlled clinical trail. *Health Economics* 1994; 333-345.
65. SAS Institute Staff. *SAS/STAT User's guide Version 8*. New York: Fatbrain. 2000.
66. Sackett DL & Torrance GW. The utility of different health states as perceived by the general public. *Journal of Chronic Disorders* 1978; 31: 697 -705.
67. Schoemaker PJH. The expected utility model: Its variants, purposes, evidence and limitations. *Journal of Economic Literature* 1982; 20: 529-563.
68. Selai C & Rosser R. Eliciting EuroQol descriptive data and utility scale values from patients. A feasibility study. *PharmacoEconomics* 1995; 8: 147-158.
69. Sintonen H. *The 15-D measure of health-related quality of life. II Feasibility, reliability and validity of its valuation system*. National Centre for Health Program Evaluation, Working Paper 42, Melbourne. 1994.
70. Stiggelbout AM, Kiebert GM, Kievit J, Leer JW, Habbeman JD, De Haes JC. The "utility" of the time trade-off method in cancer patients: Feasibility and proportional trade-off. *Journal of clinical Epidemiology* 1995; 48: 1207-1214.
71. Stiggelbout AM, Eijkemans MJC, Kiebert GM, Kieveit J, Leer JWH, De Haes HJCJM. The 'utility' of the visual analog scale in medical decision making and technology assessment: Is it an alternative to the time trade-off? *Journal of Technology Assessment in Health Care* 1996; 12: 291-298.
72. Sutherland HJ, Llewellyn-Thomas H, Boyd NF, Till JE. Attitudes towards quality of survival: The concept of maximum endurable time. *Medical Decision Making* 1982; 2: 299-309.
73. Torrance GW. Social preferences for health states: An empirical evaluation of three measurement techniques. *Socio-Economic Planning Science* 1976; 10: 129-36.
74. Torrance GW, Boyle MH, Horwood SP. Application of multi-attribute utility theory to measure social preferences for health states. *Operations Research* 1982; 30: 1043-1069.
75. Torrance GW. Measurement of health state utilities for economic appraisal. A review. *Journal of Health Economics* 1986; 5: 1-30.
76. Torrance GW, Feeny DH, Furlong WJ, Barr RD, Zhang Y, Wang Q. Multiattribute utility function for a comprehensive health status classification system. Health Utility Index Mark 2. *Medical Care* 1996; 34: 702-722.
77. Thomee R, Grimby G, Wright BD, Linacre JM. Rasch analysis of visual analog scale measurements before and after treatment of patellofemoral pain syndrome in women. *Scandinavian Journal of Rehabilitation Medicine* 1995; 27: 145-151.
78. Tsuchiya A, Ikeda S, Ikegami N et al. Estimating an EQ-5D population value set: the case of Japan. *Health Economics* 2002; 11: 341-353.
79. Weijnen T, Nieuwenhuizen M, Oppe S, de Charro F. *A European EQ-5D value set. Myth or reality?* Report nr. 6 on the VAS values. The Rotterdam Analysis Team, 1999.

80. Weinstein MC, Siegel JE, Gold MR, Kamlet MS, Russell LB. Recommendations of the panel on cost-effectiveness in health and medicine. *Journal of American Medical Association* 1996; 276: 1253-1258.

81. Williams A. Is the QALY a technical solution to a political problem? Of course not! *International Journal of Health Services* 1991; 21: 365-369.

82. Wolfson AD. Sinclair AJ, Bombadier C, McGreer A. Preference measurements for functional status in stroke patients: Inter-rater and inter-technique comparisons. In Kane RL & Kane RA (eds.) *Values and long-term care*. Lexicon Books: New York. 1982.

83. Wright BD & Linacre JM. Observations are always ordinal; measurements, however, must be interval. *Archives of Physical and Medical Rehabilitation* 1989; 70: 857-860.

Appendix A

Table A. Estimates of biased models including all inconsistencies (n =1,332).

Model / Variable	1	2	5	6	7	8	9	10	11
a	-0.0734	-0.05373	0.0072	0.0360	0.3868	1.4175	0.0911	0.0798	0.1383
	(<0.0001)	(<0.0001)	(0.5867)	(0.0070)	(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)
MO	0.1823	0.1776	0.1085	-0.0039		-0.1506	0.1059	0.0524	0.1988
	(<0.0001)	(<0.0001)	(0.0001)	(0.7523)		(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)
SC	0.1258	0.1104	0.0717	0.0156		-0.2293	0.0331	0.0203	0.0908
	(<0.0001)	(<0.0001)	(0.0001)	(0.2186)		(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)
UA	0.0709	-0.0135	0.0727	0.0082		-0.3255	-0.0577	0.0767	0.0010
	(<0.0001)	(<0.0118)	(0.0001)	(0.4460)		(<0.0001)	(<0.0001)	(0.0215)	(0.8583)
PD	0.1858	0.1427	0.1162	0.1843		-0.1694	0.0991	0.0590	0.1257
	(<0.0001)	(<0.0001)	(0.0001)	(<0.0001)		(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)
AD	0.1885	0.1369	0.1740	0.2033		-0.1780	0.0874	0.2992	0.1244
	(<0.0001)	(<0.0001)	(0.0001)	(<0.0001)		(<0.0001)	(<0.0001)	(<0.0001)	(<0.0001)
M2			0.3162					0.0511	
			(0.0001)					(0.0003)	
S2			0.1209					0.0141	
			(0.0001)					(0.3317)	
U2			-0.0446					0.1951	
			(0.0303)					(<0.0001)	
P2			0.3612					0.2040	
			(0.0001)					(<0.0001)	
A2			0.2716						
			(0.0001)						
MOSC			0.0358	0.0633					
			(0.0019)	(<0.0001)					
MOUA			-0.0236	0.1048					
			(0.0416)	(<0.0001)					
MOPD			-0.0453	0.0519					
			(0.0001)	(<0.0001)					
MOAD			-0.0227	-0.0070					
			(0.0031)	(0.3769)					
SCUA			0.0182	0.0656					
			(0.0669)	(<0.0001)					
SCPD			-0.0262	-0.0509					
			(0.0036)	(<0.0001)					
SCAD			-0.0581	-0.0008					
			(0.0001)	(0.9223)					
UAPD			0.0254	-0.0433					
			(0.0031)	(<0.0001)					
UAAD			0.0097	-0.0288					
			(0.2871)	(0.0004)					
PDAD			-0.0626	0.0113					

			(0.0001)	(0.0606)					
F11					-0.0974 (<0.0001)	-0.2635 (<0.0001)			
F21					-0.1390 (<0.0001)	-0.5741 (<0.0001)			
F31					-0.2359 (<0.0001)	-0.7979 (<0.0001)			
F41					-0.2081 (<0.0001)	-1.0521 (<0.0001)			
F13					0.2959 (<0.0001)	0.5614 (<0.0001)	0.2882 (<0.0001)		
F23					0.5647 (<0.0001)	0.9887 (<0.0001)	0.4575 (<0.0001)		
F33					0.7866 (<0.0001)	1.4711 (<0.0001)	0.6864 (<0.0001)		
F43					0 .	0 .	0 .		
F53					1.1139 (<0.0001)	2.1885 (<0.0001)	0.8740 (<0.0001)		
N2								0.0112 (0.3624)	-0.1855 (<0.0001)
N3		0.3044 (<0.0001)						0.1584 (<0.0001)	0.2633 (<0.0001)
R2	0.6078	0.6326	0.6635	0.6284	0.6383	0.6551	0.6547	0.6491	0.6322

Appendix B

Table B. Comparison of rank ordering and corresponding EQ-5D tariffs across countries; Denmark, the UK, Japan and Spain. Differences in ranking position from Denmark in square brackets.

EQ-5D health states	Denmark		UK		Japan		Spain	
	Ranking	Tariff	Ranking	Tariff	Ranking	Tariff	Ranking	Tariff
11111	1	1.000	1 [0]	1.000	1 [0]	1.000	1 [0]	1.000
11112	4	0.853	4 [0]	0.848	4 [0]	0.790	2 [-2]	0.914
11113	65	0.434	40 [-25]	0.414	12 [-53]	0.730	32 [-33]	0.541
11121	6	0.836	8 [2]	0.796	6 [0]	0.769	4 [-2]	0.887
11122	15	0.777	16 [1]	0.725	18 [3]	0.707	8 [-7]	0.825
11123	79	0.358	63 [-16]	0.291	38 [-41]	0.647	37 [-42]	0.452
11131	69	0.408	69 [0]	0.264	36 [-33]	0.649	40 [-29]	0.424
11132	83	0.349	94 [11]	0.193	64 [-19]	0.587	47 [-36]	0.362
11133	172	0.089	150 [-22]	0.028	85 [-87]	0.541	65 [-107]	0.280
11211	2	0.890	2 [0]	0.883	2 [0]	0.812	3 [1]	0.905
11212	8	0.831	7 [-1]	0.812	8 [0]	0.750	6 [-2]	0.843
11213	68	0.412	44 [-24]	0.378	22 [-46]	0.690	36 [-32]	0.470
11221	10	0.814	11 [1]	0.760	13 [3]	0.729	9 [-1]	0.816
11222	21	0.755	21 [0]	0.689	28 [7]	0.667	15 [-6]	0.754
11223	87	0.336	73 [-14]	0.255	53 [-34]	0.607	44 [-43]	0.381
11231	72	0.386	81 [9]	0.228	51 [-21]	0.609	49 [-23]	0.353
11232	90	0.327	106 [16]	0.157	82 [-8]	0.547	62 [-28]	0.291
11233	176	0.067	164 [-12]	-0.008	105 [-71]	0.501	86 [-90]	0.209
11311	29	0.698	31 [2]	0.556	17 [-12]	0.710	35 [6]	0.490
11312	34	0.639	35 [1]	0.485	37 [3]	0.648	39 [5]	0.428
11313	75	0.379	55 [-20]	0.320	54 [-21]	0.602	51 [-24]	0.346
11321	37	0.622	38 [1]	0.433	45 [8]	0.627	42 [5]	0.401
11322	43	0.563	48 [5]	0.362	72 [29]	0.565	52 [9]	0.339
11323	96	0.303	91 [-5]	0.197	95 [-1]	0.519	69 [-27]	0.257
11331	81	0.353	101 [20]	0.170	94 [13]	0.521	79 [-2]	0.229
11332	99	0.294	127 [28]	0.099	124 [25]	0.459	98 [-1]	0.167
11333	180	0.034	181 [1]	-0.066	141 [-39]	0.413	128 [-52]	0.085
12111	5	0.846	5 [0]	0.815	3 [-2]	0.799	7 [2]	0.842
12112	13	0.787	13 [0]	0.744	10 [-3]	0.737	13 [0]	0.780
12113	77	0.368	58 [-19]	0.310	24 [-53]	0.677	41 [-36]	0.407
12121	17	0.770	19 [2]	0.692	15 [-2]	0.716	16 [-1]	0.753
12122	26	0.711	27 [1]	0.621	33 [7]	0.654	22 [-4]	0.691
12123	102	0.292	96 [-6]	0.187	60 [-42]	0.594	55 [-47]	0.318
12131	86	0.342	104 [18]	0.160	58 [-28]	0.596	63 [-23]	0.290
12132	107	0.283	131 [24]	0.089	88 [-19]	0.534	80 [-27]	0.228
12133	184	0.023	184 [0]	-0.076	109 [-75]	0.488	105 [-79]	0.146
12211	9	0.824	10 [1]	0.779	7 [-2]	0.759	14 [5]	0.771
12212	19	0.765	18 [-1]	0.708	20 [1]	0.697	21 [2]	0.709
12213	85	0.346	65 [-20]	0.274	40 [-45]	0.637	53 [-32]	0.336
12221	22	0.748	24 [2]	0.656	25 [3]	0.676	23 [1]	0.682
12222	31	0.689	30 [-1]	0.585	49 [18]	0.614	28 [-3]	0.620
12223	112	0.270	107 [-5]	0.151	76 [-36]	0.554	73 [-39]	0.247
12231	93	0.320	118 [25]	0.124	75 [-18]	0.556	82 [-11]	0.219
12232	115	0.261	142 [27]	0.053	108 [-7]	0.494	102 [-13]	0.157
12233	189	0.001	192 [3]	-0.112	127 [-62]	0.448	129 [-60]	0.075
12311	36	0.632	36 [0]	0.452	31 [-5]	0.657	48 [12]	0.356
12312	41	0.573	43 [2]	0.381	59 [18]	0.595	61 [20]	0.294
12313	94	0.313	84 [-10]	0.216	80 [-14]	0.549	84 [-10]	0.212
12321	44	0.556	53 [9]	0.329	68 [24]	0.574	68 [24]	0.267
12322	53	0.497	72 [19]	0.258	100 [47]	0.512	88 [35]	0.205
12323	122	0.237	129 [7]	0.093	121[-1]	0.466	110 [-12]	0.123
12331	105	0.287	139 [34]	0.066	120 [15]	0.468	125 [20]	0.095

12332	128	0.228	163 [35]	-0.005	144 []16	0.406	145 [17]	0.033
12333	194	-0.032	208 [14]	-0.170	155 [-39]	0.360	167 [-27]	-0.049
13111	40	0.574	37 [-3]	0.436	9 [-31]	0.737	45 [5]	0.376
13112	49	0.515	46 [-3]	0.365	26 [-23]	0.675	57 [8]	0.314
13113	118	0.255	90 [-28]	0.200	43 [-75]	0.629	34 [-84]	0.509
13121	52	0.498	56 [4]	0.313	32 [-20]	0.654	64 [12]	0.287
13122	63	0.439	76 [13]	0.242	61 [-2]	0.592	81 [18]	0.225
13123	144	0.179	136 [-8]	0.077	84 [-60]	0.546	106 [-38]	0.143
13131	127	0.229	144 [17]	0.050	81 [-46]	0.548	117 [-10]	0.115
13132	149	0.170	168 [19]	-0.021	111 [-38]	0.486	136 [-13]	0.053
13133	207	-0.090	211 [4]	-0.186	131 [-76]	0.440	161 [-46]	-0.029
13211	45	0.552	41 [-4]	0.400	19 [-26]	0.697	58 [13]	0.305
13212	55	0.493	54 [-1]	0.329	41 [-14]	0.635	75 [20]	0.243
13213	125	0.233	102 [-23]	0.164	62 [-63]	0.589	100 [-25]	0.161
13221	56	0.476	64 [8]	0.277	48 [-8]	0.614	83 [27]	0.216
13222	67	0.417	87 [20]	0.206	77 [10]	0.552	103 [36]	0.154
13223	153	0.157	146 [-7]	0.041	104 [-49]	0.506	131 [-22]	0.072
13231	135	0.207	156 [21]	0.014	103 [-32]	0.508	140 [5]	0.044
13232	156	0.148	179 [23]	-0.057	128 [-28]	0.446	158 [2]	-0.018
13233	211	-0.112	216 [5]	-0.222	145 [-66]	0.400	183 [-28]	-0.100
13311	46	0.519	50 [4]	0.342	52 [6]	0.609	95 [49]	0.181
13312	58	0.460	67 [9]	0.271	83 [25]	0.547	115 [57]	0.119
13313	139	0.200	123 [-16]	0.106	106 [-33]	0.501	143 [4]	0.037
13321	60	0.443	83 [23]	0.219	92 [32]	0.526	127 [67]	0.092
13322	73	0.384	109 [36]	0.148	122 [49]	0.464	148 [75]	0.030
13323	161	0.124	166 [5]	-0.017	139 [-22]	0.418	169 [8]	-0.052
13331	146	0.174	175 [29]	-0.044	137 [-9]	0.420	179 [33]	-0.080
13332	164	0.115	195 [31]	-0.115	156 [-8]	0.358	193 [29]	-0.142
13333	217	-0.145	225 [8]	-0.280	167 [-50]	0.312	212 [-5]	-0.224
21111	3	0.857	3 [0]	0.850	5 [2]	0.774	5 [2]	0.870
21112	11	0.798	9 [-2]	0.779	16 [5]	0.712	10 [-1]	0.808
21113	76	0.379	49 [-27]	0.345	34 [-42]	0.652	38 [-38]	0.435
21121	14	0.781	15 [1]	0.727	21 [7]	0.691	12 [-2]	0.781
21122	24	0.722	23 [-1]	0.656	44 [20]	0.629	19 [-5]	0.719
21123	97	0.303	82 [-15]	0.222	71 [-26]	0.569	50 [-47]	0.346
21131	82	0.353	93 [11]	0.195	69 [-13]	0.571	56 [-26]	0.318
21132	100	0.294	117 [17]	0.124	102 [2]	0.509	70 [-30]	0.256
21133	181	0.034	173 [-8]	-0.041	123 [-58]	0.463	97 [-84]	0.174
21211	7	0.835	6 [-1]	0.814	11 [4]	0.734	11 [4]	0.799
21212	16	0.776	14 [-2]	0.743	27 [11]	0.672	17 [1]	0.737
21213	80	0.357	59 [-21]	0.309	50 [-30]	0.612	46 [-34]	0.364
21221	20	0.759	20 [0]	0.691	35 [15]	0.651	20 [0]	0.710
21222	28	0.700	28 [0]	0.620	63 [35]	0.589	26 [-2]	0.648
21223	109	0.281	97 [-12]	0.186	91 [-18]	0.529	66 [-43]	0.275
21231	88	0.331	105 [17]	0.159	90 [2]	0.531	74 [-14]	0.247
21232	110	0.272	132 [22]	0.088	118 [8]	0.469	92 [-18]	0.185
21233	185	0.012	185 [0]	-0.077	136 [-49]	0.423	123 [-62]	0.103
21311	33	0.643	34 [1]	0.487	42 [9]	0.632	43 [10]	0.384
21312	38	0.584	39 [1]	0.416	70 [32]	0.570	54 [16]	0.322
21313	92	0.324	74 [-18]	0.251	93 [1]	0.524	76 [-16]	0.240
21321	42	0.567	47 [5]	0.364	79 [37]	0.549	60 [18]	0.295
21322	50	0.508	62 [12]	0.293	110 [60]	0.487	77 [27]	0.233
21323	120	0.248	115 [-5]	0.128	130 [10]	0.441	104 [-16]	0.151
21331	98	0.298	125 [27]	0.101	129 [31]	0.443	111 [13]	0.123
21332	121	0.239	149 [28]	0.030	149 [28]	0.381	132 [11]	0.061
21333	192	-0.021	200 [8]	-0.135	161 [-31]	0.335	159 [-33]	-0.021
22111	12	0.791	12 [0]	0.746	14 [2]	0.721	18 [6]	0.736
22112	23	0.732	22 [-1]	0.675	29 [6]	0.659	24 [1]	0.674
22113	95	0.313	78 [-17]	0.241	55 [-40]	0.599	59 [-36]	0.301
22121	25	0.715	26 [1]	0.623	39 [14]	0.638	27 [2]	0.647
22122	32	0.656	32 [0]	0.552	66 [34]	0.576	30 [-2]	0.585
22123	123	0.237	121 [-2]	0.118	98 [-25]	0.516	85 [-38]	0.212
22131	106	0.287	130 [24]	0.091	96 [-10]	0.518	94 [-12]	0.184
22132	129	0.228	153 [24]	0.020	125 [-4]	0.456	113 [-16]	0.122

22133	195	-0.032	202 [7]	-0.145	142 [-53]	0.410	142 [-53]	0.040
22211	18	0.769	17 [-1]	0.710	23 [5]	0.681	25 [7]	0.665
22212	27	0.710	25 [-2]	0.639	46 [19]	0.619	29 [2]	0.603
22213	103	0.291	88 [-15]	0.205	73 [-30]	0.559	78 [-25]	0.230
22221	30	0.693	29 [-1]	0.587	56 [26]	0.598	31 [1]	0.576
22222	35	0.634	33 [-2]	0.516	86 [51]	0.536	33 [-2]	0.514
22223	133	0.215	134 [1]	0.082	113 [-20]	0.476	107 [-26]	0.141
22231	114	0.265	141 [27]	0.055	112 [-2]	0.478	118 [4]	0.113
22232	137	0.206	165 [28]	-0.016	140 [3]	0.416	137 [0]	0.051
22233	199	-0.054	209 [10]	-0.181	151 [-48]	0.370	164 [-35]	-0.031
22311	39	0.577	42 [3]	0.383	65 [26]	0.579	72 [33]	0.250
22312	48	0.518	57 [9]	0.312	97 [49]	0.517	91 [43]	0.188
22313	117	0.258	110 [-7]	0.147	115 [-2]	0.471	121 [4]	0.106
22321	51	0.501	70 [19]	0.260	107 [56]	0.496	101 [50]	0.161
22322	62	0.442	95 [33]	0.189	132 [70]	0.434	124 [62]	0.099
22323	143	0.182	152 [9]	0.024	147 [4]	0.388	150 [7]	0.017
22331	126	0.232	161 [35]	-0.003	146 [20]	0.390	155 [29]	-0.011
22332	148	0.173	183 [35]	-0.074	162 [14]	0.328	175 [27]	-0.073
22333	206	-0.087	219 [13]	-0.239	173 [-33]	0.282	196 [-10]	-0.155
23111	47	0.519	45 [-2]	0.367	30 [-17]	0.659	67 [20]	0.270
23112	59	0.460	61[2]	0.296	57 [-2]	0.597	87 [28]	0.208
23113	140	0.200	114 [-26]	0.131	78 [-62]	0.551	109 [-31]	0.126
23121	61	0.443	75 [14]	0.244	67 [6]	0.576	96 [35]	0.181
23122	74	0.384	99 [25]	0.173	99 [25]	0.514	116 [42]	0.119
23123	162	0.124	157 [-5]	0.008	119 [-43]	0.468	144 [-18]	0.037
23131	147	0.174	167 [20]	-0.019	116 [-31]	0.470	152 [5]	0.009
23132	165	0.115	188 [23]	-0.090	143 [-22]	0.408	170 [5]	-0.053
23133	218	-0.145	221 [3]	-0.255	154 [-64]	0.362	189 [-29]	-0.135
23211	54	0.497	52 [-2]	0.331	47 [-7]	0.619	89 [35]	0.199
23212	64	0.438	71 [7]	0.260	74 [10]	0.557	108 [44]	0.137
23213	145	0.178	128 [-17]	0.095	101 [-44]	0.511	135 [-10]	0.055
23221	66	0.421	86 [20]	0.208	87 [21]	0.536	120 [54]	0.110
23222	78	0.362	113 [35]	0.137	114 [36]	0.474	139 [61]	0.048
23223	168	0.102	170 [2]	-0.028	134 [-34]	0.428	165 [-3]	-0.034
23231	154	0.152	177 [23]	-0.055	133 [-21]	0.430	172 [18]	-0.062
23232	170	0.093	198 [28]	-0.126	152 [-18]	0.368	187 [17]	-0.124
23233	220	-0.167	227 [7]	-0.291	164 [-56]	0.322	207 [-13]	-0.206
23311	57	0.464	66 [9]	0.273	89 [32]	0.531	130 [73]	0.075
23312	70	0.405	89 [19]	0.202	117 [47]	0.469	151 [81]	0.013
23313	159	0.145	148 [-11]	0.037	135 [-24]	0.423	174 [15]	-0.069
23321	71	0.388	108 [37]	0.150	126 [55]	0.448	157 [86]	-0.014
23322	89	0.329	135 [46]	0.079	148 [59]	0.386	177 [88]	-0.076
23323	175	0.069	186 [11]	-0.086	159 [-16]	0.340	197 [22]	-0.158
23331	163	0.119	193 [30]	-0.113	158 [-5]	0.342	201 [38]	-0.186
23332	177	0.060	210 [33]	-0.184	174 [-3]	0.280	214 [37]	-0.248
23333	226	-0.200	233 [7]	-0.349	185 [-41]	0.234	225 [-1]	-0.330
31111	84	0.348	51 [-33]	0.336	138 [54]	0.420	71 [-13]	0.255
31112	104	0.289	68 [-36]	0.265	157 [53]	0.358	90 [-14]	0.193
31113	183	0.029	126 [-57]	0.100	168 [-15]	0.312	119 [-64]	0.111
31121	111	0.272	85 [-26]	0.213	160 [49]	0.337	99 [-12]	0.166
31122	134	0.213	111 [-23]	0.142	176 [42]	0.275	122 [-12]	0.104
31123	197	-0.047	169 [-28]	-0.023	188 [-9]	0.229	149 [-48]	0.022
31131	188	0.003	176 [-12]	-0.050	186 [-2]	0.231	154 [-34]	-0.006
31132	200	-0.056	197 [-3]	-0.121	205 [5]	0.169	173 [-27]	-0.068
31133	235	-0.316	226 [-9]	-0.286	216 [-19]	0.123	195 [-40]	-0.150
31211	91	0.326	60 [-31]	0.300	150 [59]	0.380	93 [2]	0.184
31212	113	0.267	80 [-33]	0.229	166 [53]	0.318	112 [-1]	0.122
31213	187	0.007	140 [-47]	0.064	177 [-10]	0.272	141 [-46]	0.040
31221	119	0.250	98 [-21]	0.177	170 [51]	0.297	126 [7]	0.095
31222	141	0.191	124 [-17]	0.106	184 [43]	0.235	146 [5]	0.033
31223	203	-0.069	180 [-23]	-0.059	197 [-6]	0.189	168 [-35]	-0.049
31231	190	-0.019	187 [-3]	-0.086	195 [5]	0.191	178 [-12]	-0.077
31232	204	-0.078	205 [1]	-0.157	213 [9]	0.129	191 [-13]	-0.139
31233	236	-0.338	230 [-6]	-0.322	224 [-12]	0.083	211 [-25]	-0.221

31311	101	0.293	77 [-24]	0.242	171 [70]	0.292	133 [32]	0.060
31312	124	0.234	100 [-24]	0.171	187 [63]	0.230	153 [29]	-0.002
31313	193	-0.026	158 [-35]	0.006	198 [5]	0.184	180 [-13]	-0.084
31321	132	0.217	120 [-12]	0.119	193 [61]	0.209	162 [30]	-0.029
31322	152	0.158	145 [-7]	0.048	207 [55]	0.147	181 [29]	-0.091
31323	209	-0.102	196 [-13]	-0.117	220 [11]	0.101	199 [-10]	-0.173
31331	198	-0.052	201 [3]	-0.144	219 [21]	0.103	204 [6]	-0.201
31332	210	-0.111	214 [4]	-0.215	232 [22]	0.041	216 [6]	-0.263
31333	237	-0.371	236 [-1]	-0.380	238 [1]	-0.005	228 [-9]	-0.345
32111	108	0.282	79 [-29]	0.232	153 [45]	0.367	114 [6]	0.121
32112	131	0.223	103 [-28]	0.161	169 [38]	0.305	134 [3	0.059
32113	196	-0.037	162 [-34]	-0.004	179 [-17]	0.259	160 [-36]	-0.023
32121	136	0.206	122 [-14]	0.109	172 [36]	0.284	147 [11]	0.032
32122	157	0.147	147 [-10]	0.038	189 [32]	0.222	163 [6]	-0.030
32123	212	-0.113	199 [-13]	-0.127	202 [-10]	0.176	185 [-27]	-0.112
32131	202	-0.063	204 [2]	-0.154	200 [-2]	0.178	192 [-10]	-0.140
32132	214	-0.122	217 [3]	-0.225	217 [3]	0.116	205 [-9]	-0.202
32133	238	-0.382	237 [-1]	-0.390	227 [-11]	0.070	220 [-18]	-0.284
32211	116	0.260	92 [-24]	0.196	163 [47]	0.327	138 [22]	0.050
32212	138	0.201	116 [-22]	0.125	178 [40]	0.265	156 [18]	-0.012
32213	201	-0.059	172 [-29]	-0.040	190 [-11]	0.219	182 [-19]	-0.094
32221	142	0.184	137 [-5]	0.073	181 [39]	0.244	166 [24]	-0.039
32222	160	0.125	159 [-1]	0.002	199 [39]	0.182	184 [24]	-0.101
32223	215	-0.135	206 [-9]	-0.163	209 [-6]	0.136	200 [-15]	-0.183
32231	205	-0.085	212 [7]	-0.190	208 [3]	0.138	208 [3]	-0.211
32232	216	-0.144	222 [6]	-0.261	226 [10]	0.076	218 [2]	-0.273
32233	239	-0.404	238 [-1]	-0.426	233 [-6]	0.030	229 [-10]	-0.355
32311	130	0.227	112 [-18]	0.138	182 [52]	0.239	176 [46]	-0.074
32312	151	0.168	138 [-13]	0.067	201 [50]	0.177	190 [39]	-0.136
32313	208	-0.092	190 [-18]	-0.098	211 [3]	0.131	210 [2]	-0.218
32321	155	0.151	154 [-1]	0.015	206 [51]	0.156	198 [43]	-0.163
32322	171	0.092	178 [7]	-0.056	221 [50]	0.094	213 [42]	-0.225
32323	221	-0.168	215 [-6]	-0.221	230 [9]	0.048	222 [1]	-0.307
32331	213	-0.118	220 [7]	-0.248	229 [16]	0.050	226 [13]	-0.335
32332	224	-0.177	229 [5]	-0.319	239 [15]	-0.012	234 [10]	-0.397
32333	240	-0.437	240 [0]	-0.484	241 [1]	-0.058	238 [-2]	-0.479
33111	150	0.169	119 [-31]	0.122	165 [15]	0.319	171 [21]	-0.054
33112	167	0.110	143 [-24]	0.051	180 [13	0.257	186 [19]	-0.116
33113	219	-0.150	194 [-25]	-0.114	192 [-27]	0.211	203 [-16]	-0.198
33121	169	0.093	160 [-9]	-0.001	183 [14]	0.236	194 [25]	-0.143
33122	182	0.034	182 [0]	-0.072	203 [21]	0.174	206 [24]	-0.205
33123	228	-0.226	218 [-10]	-0.237	215 [-13]	0.128	221 [-7]	-0.287
33131	223	-0.176	223 [0]	-0.264	212 [-11]	0.130	224 [1]	-0.315
33132	230	-0.235	232 [2]	-0.335	228 [-2]	0.068	231 [1]	-0.377
33133	241	-0.495	241 [0]	-0.500	235 [-6]	0.022	237 [-4]	-0.459
33211	158	0.147	133 [-25]	0.086	175 [17]	0.279	188 [30]	-0.125
33212	173	0.088	155 [-18]	0.015	191 [18]	0.217	202 [29]	-0.187
33213	222	-0.172	203 [-19]	-0.150	204 [-18]	0.171	217 [-5]	-0.269
33221	174	0.071	171 [-3]	-0.037	194 [20]	0.196	209 [35]	-0.214
33222	186	0.012	191 [5]	-0.108	210 [24]	0.134	219 [33]	-0.276
33223	231	-0.248	224 [-7]	-0.273	223 [-8]	0.088	230 [-1]	-0.358
33231	225	-0.198	228 [3]	-0.300	222 [-3]	0.090	232 [7]	-0.386
33232	232	-0.257	235 [3]	-0.371	234 [2]	0.028	236 [4]	-0.448
33233	242	-0.517	242 [0]	-0.536	240 [-2]	-0.018	241 [-1]	-0.530
33311	166	0.114	151 [-15]	0.028	196 [30]	0.191	215 [49]	-0.249
33312	178	0.055	174 [-4]	-0.043	214 [36]	0.129	223 [45]	-0.311
33313	227	-0.205	213 [-14]	-0.208	225 [-2]	0.083	233 [6]	-0.393
33321	179	0.038	189 [10]	-0.095	218 [39]	0.108	227 [48]	-0.338
33322	191	-0.021	207 [16]	-0.166	231 [40]	0.046	235 [44]	-0.400
33323	233	-0.281	231 [-2]	-0.331	237 [4]	0.000	239 [6]	-0.482
33331	229	-0.231	234 [5]	-0.358	236 [7]	0.002	240 [11]	-0.510
33332	234	-0.290	239 [5]	-0.429	242 [8]	-0.060	242 [8]	-0.572
33333	243	-0.550	243 [0]	-0.594	243 [0]	-0.106	243 [0]	-0.654

Appendix C

The Danish EQ-5D tariffs (based on the TTO3 model)

11111	1.000	13222	0.516	22333	-0.137	32221	0.302
11112	0.818	13223	0.217	23111	0.641	32222	0.234
11113	0.519	13231	0.251	23112	0.573	32223	-0.066
11121	0.824	13232	0.183	23113	0.274	32231	-0.032
11122	0.756	13233	-0.117	23121	0.579	32232	-0.100
11123	0.456	13311	0.551	23122	0.511	32233	-0.399
11131	0.490	13312	0.483	23123	0.211	32311	0.269
11132	0.422	13313	0.183	23131	0.245	32312	0.200
11133	0.123	13321	0.489	23132	0.177	32313	-0.099
11211	0.838	13322	0.421	23133	-0.122	32321	0.206
11212	0.770	13323	0.121	23211	0.594	32322	0.138
11213	0.471	13331	0.155	23212	0.525	32323	-0.161
11221	0.776	13332	0.087	23213	0.226	32331	-0.127
11222	0.708	13333	-0.213	23221	0.531	32332	-0.196
11223	0.409	21111	0.833	23222	0.463	32333	-0.495
11231	0.442	21112	0.765	23223	0.164	33111	0.283
11232	0.374	21113	0.465	23231	0.198	33112	0.215
11233	0.075	21121	0.771	23232	0.129	33113	-0.084
11311	0.743	21122	0.703	23233	-0.170	33121	0.221
11312	0.674	21123	0.403	23311	0.498	33122	0.153
11313	0.375	21131	0.437	23312	0.430	33123	-0.146
11321	0.680	21132	0.369	23313	0.130	33131	-0.113
11322	0.612	21133	0.069	23321	0.436	33132	-0.181
11323	0.313	21211	0.785	23322	0.367	33133	-0.480
11331	0.347	21212	0.717	23323	0.068	33211	0.236
11332	0.278	21213	0.418	23331	0.102	33212	0.167
11333	-0.021	21221	0.723	23332	0.034	33213	-0.132
12111	0.823	21222	0.655	23333	-0.266	33221	0.173
12112	0.755	21223	0.355	31111	0.475	33222	0.105
12113	0.456	21231	0.389	31112	0.407	33223	-0.194
12121	0.761	21232	0.321	31113	0.107	33231	-0.160
12122	0.693	21233	0.021	31121	0.413	33232	-0.229
12123	0.393	21311	0.689	31122	0.345	33233	-0.528
12131	0.427	21312	0.621	31123	0.045	33311	0.140
12132	0.359	21313	0.322	31131	0.079	33312	0.072
12133	0.060	21321	0.627	31132	0.011	33313	-0.228
12211	0.776	21322	0.559	31133	-0.289	33321	0.078
12212	0.707	21323	0.260	31211	0.427	33322	0.009
12213	0.408	21331	0.293	31212	0.359	33323	-0.290
12221	0.713	21332	0.225	31213	0.060	33331	-0.256
12222	0.645	21333	-0.074	31221	0.365	33332	-0.324
12223	0.346	22111	0.770	31222	0.297	33333	-0.624
12231	0.380	22112	0.702	31223	-0.003		
12232	0.311	22113	0.402	31231	0.031	[Dead]	[0.000]
12233	0.012	22121	0.708	31232	-0.037	[UNC]	[-0.293]
12311	0.680	22122	0.640	31233	-0.336		
12312	0.612	22123	0.340	31311	0.331		
12313	0.312	22131	0.374	31312	0.263		
12321	0.618	22132	0.306	31313	-0.036		
12322	0.549	22133	0.006	31321	0.269		
12323	0.250	22211	0.722	31322	0.201		
12331	0.284	22212	0.654	31323	-0.098		
12332	0.216	22213	0.355	31331	-0.065		
12333	-0.084	22221	0.660	31332	-0.133		
13111	0.695	22222	0.592	31333	-0.432		
13112	0.626	22223	0.292	32111	0.412		
13113	0.327	22231	0.326	32112	0.344		
13121	0.632	22232	0.258	32113	0.044		
13122	0.564	22233	-0.041	32121	0.350		
13123	0.265	22311	0.627	32122	0.282		
13131	0.299	22312	0.558	32123	-0.018		
13132	0.230	22313	0.259	32131	0.016		
13133	-0.069	22321	0.564	32132	-0.052		
13211	0.647	22322	0.496	32133	-0.352		
13212	0.579	22323	0.197	32211	0.364		
13213	0.279	22331	0.231	32212	0.296		
13221	0.585	22332	0.162	32213	-0.003		

Appendix D

Table D. Differences between Danish TTO-based and VAS-based EQ-5D tariffs. Same data set.

Health states	TTO tariff	VAS tariff	Differences	Health states	TTO tariff	VAS tariff	Differences
11111	1.000	1.000	0.000	12333	-0.084	0.006	-0.090
11112	0.818	0.696	0.122	13111	0.695	0.726	-0.031
11113	0.519	0.528	-0.010	13112	0.626	0.558	0.068
11121	0.824	0.724	0.100	13113	0.327	0.390	-0.063
11122	0.756	0.556	0.200	13121	0.632	0.586	0.046
11123	0.456	0.388	0.068	13122	0.564	0.418	0.146
11131	0.490	0.584	-0.094	13123	0.265	0.250	0.014
11132	0.422	0.416	0.006	13131	0.299	0.446	-0.148
11133	0.123	0.249	-0.126	13132	0.230	0.278	-0.048
11211	0.838	0.777	0.062	13133	-0.069	0.111	-0.180
11212	0.770	0.609	0.161	13211	0.647	0.639	0.008
11213	0.471	0.441	0.029	13212	0.579	0.471	0.108
11221	0.776	0.637	0.139	13213	0.279	0.303	-0.024
11222	0.708	0.469	0.239	13221	0.585	0.499	0.086
11223	0.409	0.301	0.107	13222	0.516	0.331	0.185
11231	0.442	0.497	-0.055	13223	0.217	0.163	0.053
11232	0.374	0.329	0.045	13231	0.251	0.359	-0.108
11233	0.075	0.162	-0.087	13232	0.183	0.191	-0.009
11311	0.743	0.690	0.053	13233	-0.117	0.024	-0.141
11312	0.674	0.522	0.152	13311	0.551	0.552	-0.001
11313	0.375	0.354	0.021	13312	0.483	0.384	0.099
11321	0.680	0.550	0.130	13313	0.183	0.216	-0.033
11322	0.612	0.382	0.230	13321	0.489	0.412	0.077
11323	0.313	0.214	0.098	13322	0.421	0.244	0.176
11331	0.347	0.410	-0.064	13323	0.121	0.076	0.045
11332	0.278	0.242	0.036	13331	0.155	0.272	-0.117
11333	-0.021	0.075	-0.096	13332	0.087	0.104	-0.018
12111	0.823	0.795	0.029	13333	-0.213	-0.063	-0.149
12112	0.755	0.627	0.128	21111	0.833	0.690	0.143
12113	0.456	0.459	-0.004	21112	0.765	0.522	0.243
12121	0.761	0.655	0.106	21113	0.465	0.355	0.111
12122	0.693	0.487	0.206	21121	0.771	0.550	0.221
12123	0.393	0.319	0.074	21122	0.703	0.383	0.320
12131	0.427	0.515	-0.088	21123	0.403	0.215	0.188
12132	0.359	0.347	0.012	21131	0.437	0.411	0.026
12133	0.060	0.180	-0.120	21132	0.369	0.243	0.126
12211	0.776	0.708	0.068	21133	0.069	0.075	-0.006
12212	0.707	0.540	0.167	21211	0.785	0.603	0.182
12213	0.408	0.372	0.036	21212	0.717	0.435	0.282
12221	0.713	0.568	0.145	21213	0.418	0.268	0.150
12222	0.645	0.400	0.245	21221	0.723	0.463	0.260
12223	0.346	0.232	0.113	21222	0.655	0.296	0.359
12231	0.380	0.428	-0.049	21223	0.355	0.128	0.228
12232	0.311	0.260	0.051	21231	0.389	0.324	0.066
12233	0.012	0.093	-0.081	21232	0.321	0.156	0.165
12311	0.680	0.621	0.059	21233	0.021	-0.012	0.033
12312	0.612	0.453	0.159	21311	0.689	0.516	0.173
12313	0.312	0.285	0.027	21312	0.621	0.348	0.273
12321	0.618	0.481	0.137	21313	0.322	0.181	0.141
12322	0.549	0.313	0.236	21321	0.627	0.376	0.251
12323	0.250	0.145	0.104	21322	0.559	0.209	0.350
12331	0.284	0.341	-0.058	21323	0.260	0.041	0.219
12332	0.216	0.173	0.042	21331	0.293	0.237	0.057
21332	0.225	0.069	0.156	31122	0.345	0.209	0.136
21333	-0.074	-0.099	0.025	31123	0.045	0.041	0.004
22111	0.770	0.738	0.032	31221	0.365	0.290	0.075
22112	0.702	0.570	0.132	31222	0.297	0.122	0.175
22113	0.402	0.403	0.000	31223	-0.003	-0.046	0.043

22121	0.708	0.598	0.110	31231	0.031	0.150	-0.119
22122	0.640	0.431	0.209	31232	-0.037	-0.018	-0.019
22123	0.340	0.263	0.078	31233	-0.336	-0.186	-0.151
22131	0.374	0.459	-0.084	31311	0.331	0.343	-0.011
22132	0.306	0.291	0.015	31312	0.263	0.175	0.088
22133	0.006	0.123	-0.117	31313	-0.036	0.007	-0.043
22211	0.722	0.651	0.071	31321	0.269	0.203	0.066
22212	0.654	0.483	0.171	31322	0.201	0.035	0.166
22213	0.355	0.316	0.039	31323	-0.098	-0.133	0.034
22221	0.660	0.511	0.149	31331	-0.065	0.063	-0.128
22222	0.592	0.344	0.248	31332	-0.133	-0.105	-0.028
22223	0.292	0.176	0.117	31333	-0.432	-0.273	-0.160
22231	0.326	0.372	-0.045	32111	0.412	0.682	-0.269
22232	0.258	0.204	0.054	32112	0.344	0.514	-0.170
22233	-0.041	0.036	-0.077	32113	0.044	0.346	-0.301
22311	0.627	0.564	0.062	32121	0.350	0.542	-0.192
22312	0.558	0.396	0.162	32122	0.282	0.374	-0.092
22313	0.259	0.229	0.030	32123	-0.018	0.206	-0.224
22321	0.564	0.424	0.140	32131	0.016	0.402	-0.386
22322	0.496	0.257	0.240	32132	-0.052	0.234	-0.286
22323	0.197	0.089	0.108	32133	-0.352	0.066	-0.418
22331	0.231	0.285	-0.054	32211	0.364	0.595	-0.230
22332	0.162	0.117	0.046	32212	0.296	0.427	-0.131
22333	-0.137	-0.051	-0.086	32213	-0.003	0.259	-0.262
23111	0.641	0.669	-0.028	32221	0.302	0.455	-0.153
23112	0.573	0.501	0.072	32222	0.234	0.287	-0.053
23113	0.274	0.334	-0.060	32223	-0.066	0.119	-0.185
23121	0.579	0.529	0.050	32231	-0.032	0.315	-0.347
23122	0.511	0.362	0.150	32232	-0.100	0.147	-0.247
23123	0.211	0.194	0.018	32233	-0.399	-0.021	-0.379
23131	0.245	0.390	-0.144	32311	0.269	0.508	-0.239
23132	0.177	0.222	-0.045	32312	0.200	0.340	-0.139
23133	-0.122	0.054	-0.176	32313	-0.099	0.172	-0.271
23211	0.594	0.582	0.011	32321	0.206	0.368	-0.161
23212	0.525	0.414	0.111	32322	0.138	0.200	-0.062
23213	0.226	0.247	-0.021	32323	-0.161	0.032	-0.194
23221	0.531	0.442	0.089	32331	-0.127	0.228	-0.355
23222	0.463	0.275	0.189	32332	-0.196	0.060	-0.256
23223	0.164	0.107	0.057	32333	-0.495	-0.108	-0.388
23231	0.198	0.303	-0.105	33111	0.283	0.613	-0.329
23232	0.129	0.135	-0.005	33112	0.215	0.445	-0.229
23233	-0.170	-0.033	-0.137	33113	-0.084	0.277	-0.361
23311	0.498	0.495	0.003	33121	0.221	0.473	-0.251
23312	0.430	0.327	0.102	33122	0.153	0.305	-0.152
23313	0.130	0.160	-0.030	33123	-0.146	0.137	-0.284
23321	0.436	0.355	0.080	33131	-0.113	0.333	-0.446
23322	0.367	0.188	0.180	33132	-0.181	0.165	-0.346
23323	0.068	0.020	0.048	33133	-0.480	-0.003	-0.478
23331	0.102	0.216	-0.114	33211	0.236	0.526	-0.290
23332	0.034	0.048	-0.014	33212	0.167	0.358	-0.190
23333	-0.266	-0.120	-0.146	33213	-0.132	0.190	-0.322
31111	0.475	0.517	-0.041	33221	0.173	0.386	-0.212
31112	0.407	0.349	0.058	33222	0.105	0.218	-0.113
31113	0.107	0.181	-0.074	33223	-0.194	0.050	-0.244
31121	0.413	0.377	0.036	33231	-0.160	0.246	-0.406
31131	0.079	0.237	-0.158	33232	-0.229	0.078	-0.307
31132	0.011	0.069	-0.058	33233	-0.528	-0.090	-0.439
31133	-0.289	-0.099	-0.190	33311	0.140	0.439	-0.299
31211	0.427	0.430	-0.002	33312	0.072	0.271	-0.199
31212	0.359	0.262	0.097	33313	-0.228	0.103	-0.331
31213	0.060	0.094	-0.034	33321	0.078	0.299	-0.221
33322	0.009	0.131	-0.122				
33323	-0.290	-0.037	-0.253				
33331	-0.256	0.159	-0.415				

33332	-0.324	-0.009	-0.316				
33333	-0.624	-0.177	-0.447				
Dead	[0.000]	[0.000]	[0.000]				
Uncon.	[-0.293]	[-0.131]	[-0.162]				