HEDS Discussion Paper 06/01


ScHARR
School of Health and
Related Research

The University of Sheffield
ScHARR
School of Health and Related Research

Health Economics and Decision Science

Discussion Paper Series

February 2006

Ref: 06/1

Modelling health state preference data using a non-parametric Bayesian method

Samer Kharroubi and Christopher McCabe

Health Economics and Decision Science

Corresponding Author:
Christopher McCabe
Health Economics and Decision Science
School of Health and Related Research
University of Sheffield
Regent Court, Sheffield, UK
S1 4DA
Email: c.mccabe@sheffield.ac.uk

Abstract

This paper reports on the findings from the application of a recently reported approach to modelling health state valuation data. The approach applies a nonparametric model to estimate the revised version of the Health Utilities Index Mark 2 (HUI 2) health state valuation algorithm using Bayesian methods. The data set is the UK HUI 2 valuation study where a sample of 51 states defined by the HUI 2 was valued by a sample of the UK general population using standard gamble. The paper presents the results from applying the nonparametric model and compares these to the original model estimated using a conventional parametric random effects model. The two models are compared in terms of their predictive performance. The paper discusses the implications of these results for future applications of the HUI 2 and further work in this field.

JEL classification: I1

Key words: Preference-based health measure; HUI 2; nonparametric methods

1. Introduction

There has been an increasing use of preference-based measures of health related quality of life in order to calculate quality adjusted life years for use in cost effectiveness analyses. These preference-based measures are standardised multi-dimensional health state classifications with pre-existing preference or utility weights elicited from a sample of the general population. There are currently a number of such preference-based measures, including the EQ-5D (Brooks, 1996), HUI2 & 3 (Torrance et al, 1996; Feeny et al, 2002), 15D (Sintonen, 1994, 1995), AQoL (Hawthorn et al, 2001), QWB (Kaplan et al, 1988) and the SF-6D (Brazier et al, 2002). All of these measures describe a large number of health states, such that it is not feasible to obtain direct valuations for each health state. Thus models are constructed to predict the values for all states in a descriptive system based upon direct valuations of a sample of states.

Health state values present a significant challenge for conventional statistical modelling procedures due to their nature, namely: skewed, truncated, non-continuous and hierarchical (Brazier et al, 2002). Attempts to statistically model these data have met with some success in the EQ-5D, SF-6D and HUI2 (Dolan, 1997 and Brazier et al, 2002, McCabe et al 2005a). However, there were concerns with the size of the prediction errors, non-monotonicity in some estimated health state values and an apparent systematic pattern in the prediction errors. Recently, Kharroubi et al (2005) reported an alternative, non-parametric Bayesian method for modelling health state preference data. This paper reports the application of this method to the UK HUI2 valuation data reported by McCabe et al, and compares the results with the conventional random effect regression model.

The next section of this paper provides a brief description of the UK HUI2 valuation survey. A detailed description has been reported elsewhere. (McCabe et al, 2005a) In section 3 the results from each approach are presented and compared in terms of their ability to predict actual values. We conclude with a discussion of the results and their implication for the HUI2 and future approaches to modelling health state preference data.

2. The Health Utilities Index Mark 2

The HUI2 is a preference-based multi-attribute health related quality of life instrument specifically developed for use with children.(Torrance et al, 1996) It consists of seven attributes (sensation, mobility, emotion, cognition, self care, pain and fertility), each of which has between three and five levels. The levels describe a range, from 'normal functioning for age' to 'extreme disability'. (Table 1). When it is used as a generic instrument, fertility is excluded. (Torrance et al, 1996). The generic version of the instrument was used for the UK valuation survey.

As part of a large study of outcomes following paediatric intensive care in the United Kingdom, 3 separate valuation surveys were undertaken. These have been described in detail elsewhere. (McCabe 2005a, b). The work reported in this paper utilises the data from two of those surveys.

One hundred and ninety nine respondents provided valuations for 51 health states in the HUI2 descriptive system, using the standard gamble technique following the methods described by Furlong et al. (1990) The mean number of valuations per health state was 24 (range 9 to 29).

Fifty one respondents provided valuations for 14 health states from the HUI2 descriptive system not valued in the Valuation Survey, using the same standard gamble technique and script as in the Valuation Survey.

3. Modelling

The generic HUI2 descriptive system describes 8000 possible health states and the empirical survey could obtain valuations for only a small subset. The aim of modelling is to estimate health state utility values for all states. The utility associated with a health state is assumed to be a function of that state, hence by estimating a relationship between the descriptive system and the observed values we can infer values for all states. Valuation surveys generate data with a complex structure creating a number of problems for estimation and a variety of techniques have been used to deal with these problems. In the main these have used parametric

relationships with particular assumptions about functional form, but here we contrast this conventional parametric approach reported by McCabe et al (2005a) with a more realistic and flexible nonparametric model.

A general model for health state valuations can be described by:

$$y_{ij} = f(\mathbf{x}_{ij}, \alpha_j) + \varepsilon_{ij}, \qquad (1)$$

where, for i = 1, 2, …, n_j and j = 1,2, …, m, $\mathbf{x}_{ij}$ is the i^{th} health state valued by respondent j and the dependent variable y_{ij} is the adjusted SG score given by respondent j for that health state. The general model has two sets of independent, zero-mean, random effect terms: ε_{ij} is a random error term associated with each observation and α_j is a term to allow for individual characteristics of respondent j.

The interpretation of $f(\mathbf{x}_{ij}, \alpha_j)$ is as the true indifference SG value that respondent j has for health state $\mathbf{x}_{ij}$. The objective is to obtain a health state utility measure for the population as a whole, and this is generally taken to be the mean of the respondent-level health state utilities across the population. In order to account for different populations, it is possible to model α_j in terms of respondent-level covariates such as age, gender or socio-economic factors, but the principal objective of the HUI2 valuation study was to estimate a health state utility function for the UK population as a whole.

3.1 The parametric approach

McCabe et al (2005a) specify the following model for respondent j's health state utility:

$$f(\mathbf{x}_{ij}, \alpha_j) = \mu + \boldsymbol{\theta}'\mathbf{I}(\mathbf{x}_{ij}) + \alpha_j, \qquad (2)$$

where μ and $\boldsymbol{\theta}$ denote unknown parameters, $\mathbf{I}(\mathbf{x}_{ij})$ is a vector of dummy explanatory variables. In the simplest, no-interactions, case of this model, $\mathbf{I}(\mathbf{x}_{ij})$ is a vector of terms $I_{\delta\lambda}(\mathbf{x}_{ij})$ for each level $\lambda > 1$ of dimension δ of the HUI 2. For example, $I_{32}(\mathbf{x}_{ij})$ denotes dimension δ = 3 (emotion), level λ = 2 (Occasionally fretful, angry,

irritable, anxious depressed or suffering from "night terrors"). For any given health state $\mathbf{x}_{ij}$, $I_{\delta\lambda}(\mathbf{x}_{ij})$ is defined as:

$I_{\delta\lambda}(\mathbf{x}_{ij}) = 1$ if, for state $\mathbf{x}_{ij}$, dimension δ is at level λ.

$I_{\delta\lambda}(\mathbf{x}_{ij}) = 0$ if, for state $\mathbf{x}_{ij}$, dimension δ is not at level λ

In all, there are 21 of these terms, with level $\lambda = 1$ acting as a baseline for each dimension. Hence the intercept parameter μ represents the health state utility value for state (111111), and summing the coefficients $\theta_{\delta\lambda}$ of the 'on' dummies derives the value of any other state.

More generally, $\mathbf{I}(\mathbf{x}_{ij})$ can include additional dummy variables to account for interactions between the levels of different dimensions, and the model selected by McCabe et al (2005a) included no such interaction term.

Estimation of this random effects model is via generalised least squares or maximum likelihood. Since α_j has zero mean, the population health state utility for state $\mathbf{x}$ in this model is simply $\mu + \boldsymbol{\theta}'\mathbf{I}(\mathbf{x})$.

3.2 The nonparametric approach

Kharroubi et al. (2005) build a new Bayesian statistical nonparametric model to describe the intrinsic characteristics of individual health state valuation data that is argued to be more theoretically appropriate than previous parametric models. For respondent *j*, the health state utility of state $\mathbf{x}_{ij}$ is

$$f(\mathbf{x}_{ij}, \alpha_j) = 1 - \exp(\alpha_j)\{1 - u(\mathbf{x}_{ij})\}. \qquad (3)$$

Note that the individual respondent term α_j enters multiplicatively rather than additively as in (2). In the Kharroubi et al. (2005) model, the distribution of α_j is

$$\alpha_j \sim N(t_j'\theta, \tau^2)$$

where t_j is the vector of covariates for respondent j. Note here that **t**'s are centred to ensure that they have zero means, and hence that the value of $\exp(\alpha)$ for a typical person is 1.

The term $u(\mathbf{x})$ is the *median* health state utility of health state **x**.[1] It is treated as an unknown function and in a nonparametric framework it therefore becomes a random variable. The model for $u(\mathbf{x})$ is

$$u(\mathbf{x}) \sim \mathrm{N}(\gamma + \boldsymbol{\beta}'\mathbf{x}, \sigma^2), \qquad (4)$$

and furthermore the values of $u(\mathbf{x})$ and $u(\mathbf{x}')$ for two different states **x** and **x**′ have a correlation $c(\mathbf{x}, \mathbf{x}')$ which decreases as the distance between **x** and **x**′ increases. The effect of this is to assert that if **x** and **x**′ describe very similar health states (in the sense that their levels are the same or close in all dimensions) their utilities will be approximately the same, and so the preference function varies smoothly as the health state changes.

Note that the mean health state utility in (3) is

$$\bar{u}(\mathbf{x}) = 1 - \bar{\alpha}\{1 - u(\mathbf{x})\},$$

where $\bar{\alpha}$ is the mean value of $\exp(\alpha)$ over the whole population. This will not in general be 1, and so the population (mean) health state utility is not the same as the median health state utility $u(\mathbf{x})$. More details of the nonparametric modelling and evaluation of $\bar{\alpha}$ are given in Kharroubi et al (2005).

The models and the programs to undertake the Bayesian approach were written in Matlab and are available on our website (http://www.shef.ac.uk/chebs).

4. Results

Given the overall aim is to predict health states valuations, the best way to compare the two models is via their predictive ability. The models are compared on Figures 1 and 2, where the predicted and actual mean values for the 51 health states valued in

[1] Consider the case when there are no covariates, the distribution of α_j is normal, so it has zero median as well as zero mean, and the median of $\exp(\alpha_j)$ is therefore 1.

the survey have been plotted with health states ordered by actual health state values. Figure 1 presents the resulting predicted mean health state valuations (dotted line) for the parametric model (2), along with actual mean health state valuations (solid line). The dashed line represents the errors obtained by the difference between the two valuations. Figure 2 presents the corresponding plots for the nonparametric model. There is a very close level of agreement between the parametric and non-parametric model. The only exception to this being state (3,1,3,3,3,1). The parametric model predicts the observed mean value considerably more accurately than the non-parametric model. We consider this finding in more details below.

Table 2 shows the inference for the mean health state utility values of the 51 health states valued in the Valuation Survey and the 12 states that were valued in the Validation Survey. For each state, Table 2 reports the observed sample mean health state utility and the predicted mean and standard deviation for the population mean health state utility from both the nonparametric and parametric models. The states marked with an asterisk were not valued in the valuation survey. Across the 51 states that were used in the study, the predictive performance of the nonparametric model is better than the parametric model overall, with a root mean square error (RMSE) of 0.055 for the nonparametric model and 0.060 for the parametric model. Very few health states are valued worse than death in either model. Finally, it can be seen that the standard deviations of the predictions are larger for the parametric model.

As always, it is important to check the validity of the assumed models. Figure 3 plots a histogram of residuals across all 1370 health state valuations for the parametric model and Figure 4 plots the corresponding residuals for the nonparametric model. According to these models we would expect these to be approximately *normal*. Figures 3 and 4 broadly support this, although there is some evidence of skewness which is more obvious in Figure 3. This is not surprising, given the negative skewness in the original SG data at the individual level. However, the degree of skewness is probably not high enough to invalidate the analyses in both models, which assume normally distributed errors. An important finding of Figures 3 and 4 is that the nonparametric model clearly yields smaller residuals, and so fits the data better. The RMSE at the individual level is 0.2267 for the nonparametric model and 0.3403 for

the parametric model. However, this in part reflects fact that the histogram for the parametric model has been calculated the frequentist way, regarding the individual random effects as being just random and so part of the error. The histogram for the nonparametric model has been calculated using the Bayesian convention, in which the random effects are handled more like a fixed effect in frequentist terms.

A better test of the validity of the model is to investigate its ability to predict the values for states that have not been used in the estimation. Data relating to 10 selected health states were removed from the estimation data, and the models fitted on data for the remaining 41 states. Table 3 presents the true sample means for the 10 omitted states, together with their predicted mean and standard deviation values from the parametric and nonparametric models estimated on the reduced data set. The predictive performance of the nonparametric model is better than the parametric model overall, with RMSEs of 0.050 and 0.090 respectively. It can be seen that the nonparametric model predicts the omitted data quite well, and better than the parametric model. It is to be noted that the predictive standard deviations here for both models are larger than those in Table 2, because the model in Table 2 is predicting the data on which it was estimated, whereas the model in Table 3 is predicting out of sample data. The parametric standard errors are larger than the nonparametric ones, primarily because the nonparametric analysis is able to make use of other evaluations by the same respondents to estimate their individual random effects, which the frequentist analysis can not do.

Figures 5 and 6 show the Q-Q plots of standardised predictive errors for the 12 health states sample means, for the parametric and nonparametric models respectively. In each figure the straight line corresponds to the theoretical $N(0,1)$ distribution. Figure 5 suggests that the parametric model is not well validated by its predictive performance. In contrast, it is apparent from Figure 6 that the nonparametric model predictions are well validated. The mean of the standardised residuals for the parametric model is -0.66 compared to -0.14 for the nonparametric one.

To capture the impact of the respondent characteristics, Figure 7 shows the histograms of the conditional posterior distribution functions of the covariates sex and

age. These results indicate that age has a strong effect, as it is not centred on zero. Sex is centred very close to zero, and thus has a negligible effect. To demonstrate the impact of adjusting for age on the mean health state values, results are presented for the states listed in Table 2 with and without adjusting for covariates. Actual UK age distribution was taken from the UK census of 2001 (www.statistics.gov.uk/**census**). These results show that the largest differences between health state values are for the most severe HUI 2 health state. Mean health state values for the pits state are -0.0256 and -0.0018 with and without adjusting for age respectively with difference of 0.0238. This difference declines as states become milder. This suggests that the magnitude of the gain of moving from a severe to mild state will be a smaller for older age groups. We consider this finding in more detail in the discussion section.

5. Discussion

This paper reports the findings from applying a new approach to modelling health state valuation data. The approach applies a nonparametric model to estimate health state utility values for the HUI2 using Bayesian methods. We have presented two sets of analyses; the first has compared a Bayesian main-effects model with the existing conventional main effects model, (McCabe 2005a). The second has utilised the flexibility of the non-parametric Bayesian method to examine the impact of covariates on health state values.

The main effects Bayesian model represents a significant improvement on the conventional model. Perhaps more importantly, the Bayesian method is more robust to reductions in the quantity of data available for model estimation, than the conventional model. The conventional model estimated on 41 health states produces a number of large prediction errors. This is perhaps unsurprising, as the Bayesian method supplements the data with prior knowledge on the correlation between health state values. It is important that the specified co-relations are acceptable to the users of the modelled values. Assuming this criterion is met, the Bayesian modelling approach makes efficient use of the information provided by typically small health state valuation surveys.

It is worth noting that the Bayesian model produces a very poor prediction for one state (313331). The explanation for this may lie in considering the nature of this health state. It combines substantial limitations in sensation, emotion, cognition, and self-care with full functioning on mobility and freedom from pain. This health state may well have been extremely difficult for respondents to visualise. Is it plausible that an individual can combine 'Able to walk, bend, lift, jump and run normally for age' with 'Requires mechanical equipment to eat, bathe, dress, or use the toilet independently'? The potential problem with the plausibility of the health state is confounded by the small number of observations for the health state; n=9. There were relatively few observations for one other health state, (122222), however both models produce reasonable predictions for this state, suggesting that the plausibility of state (313331) is the more likely explanation of the poor predictive performance. To facilitate comparison with the published statistical model we have reported the model estimated on the 51 states. (McCabe 2005a). However, Figures 8 and 9 report the results of the Bayesian main effects model estimates on the dataset with state (313331) and both states (313331) and (122222) excluded respectively. It is clearly that the Bayesian model is improved as a result of this exclusion.

As reported earlier, the covariates analysis showed that whilst gender is not an important determinant of health state values, whilst age is. For milder health states, the predicted values do not vary much with age; however, the more severe the health state, the greater the impact of age on the health state value. There are at least two possible explanations for this result. Older respondents may genuinely put a lower value on lower levels of functioning; i.e the more time you have lived, the more important it is that the remaining years are lived at a reasonable level of functional health. Alternatively, it may be that younger respondents did not adhere to the instruction that the health state would last for 60 years or assumed that over such a long period of time, advances in medical science would improve the level of health related quality of life associated with any given functional condition. It may also be that the cognitive burden of the HUI2 question was too great and the values obtained were artefacts of the process and do not reflect actual preferences.

This result is, to some degree, at odds with the existing literature on the relationship between age and health state values. Normally, it is argued that as older people have

experience of ill-health, adaptation leads them to place higher values on poor health states compared to younger people. Analyses by Kharroubi et al of the SF-6D data has found this type of relationship.(Personal Communication). Dolan's analysis of the UK EQ-5D data reported a similar relationship, however, the increasing value is only observed over a limited age range; by 50 years of age the EQ-5D data show the same decreasing value with age as seen in the HUI2 data. In addition, this relationship is more pronounced in the severe health states in both the HUI2 and EQ-5D datasets.

These results do have important implications for the use of published valuation algorithms. If the analyst is interested in obtaining an estimate of the current population mean value for a given health state, the effect of the changing age profile of the general population needs to be taken into account. (http://www.statistics.gov.uk/cci/nugget.asp?ID=949). Our results indicate that, for any given health state, the current mean values should be lower than the values of 10 years ago, because of the change in the age distribution of the population. Thus, ceteris paribus the health gain from preventing individuals entering these health states will be greater and thus more cost effective.

The observed relationship between age and utility may have implications for our understanding of the difference between patient and general population health state values. As ill-health is positively correlated with age; ceteris paribus, patients' values are likely to be lower than general population values. Thus, the degree of adaptation that leads to the observation that patients value health states more highly than the general population may be larger than currently thought. At a minimum, future analyses of comparative data should control for age when estimating the scale of divergence between patient and general population values.

The covariates model has the potential to estimate health state valuation tariffs for sub-groups of populations; and for populations with different socio-economic profiles. Such tariffs are valuable in their own right and would also be very useful in the design of HUI2 health state valuation surveys for other countries; acting as informative priors. In addition, covariates models will allow the detailed exploration of the observed international variations in the health state value models for instruments such as the SF-6D, EQ-5D and HUI2.

References

Brazier, J.E., Roberts, J. & Deverill, M., 2002. The estimation of a preference-based measure of health from the SF-36 . *Journal of Health Economics* 21, 271(292).

Brooks, R., 1996. EuroQol: the current state of play. *Health policy* 37, 53(72).

Dolan, P., 1997. Modeling valuation for Euroqol health states. *Medical Care* 35, 351(363).

Dolan, P., 2000. The effect of age on health state valuations. *J. Health Services Research and Policy* 5, 17-21.

Feeny, D. H., Furlong, W. J., Torrance, G. W., Goldsmith, C.H., Zenglong, Z., Depauw, S., Denton, M., and Boyle, M., 2002. Multi-attribute and single-attribute utility function for the Health Utility Index Mark 3 system. *Medical care* 40(20), 113(128).

Furlong, W. J., Feeny, D. H., Torrance, G. W., Bar, R., and Horsman, J. 1990. *Guide to design and development of health state utility instrumentation.* CHEPA Working Paper #90-9, McMaster University Hamilton, Ontario.

Hawthorne G., Richardson G., Atherton_Day N., 2001. A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. *Annals of Medicine* 33, 358--370.

Kaplan, R.M., & Anderson, J.P., 1988. A general health policy model: update and application. *Health Services Research 23,* 203--235.

Kharroubi, S. A., O'Hagan, A. and Brazier, J. E., 2005. Estimating Utilities from individual health state preference data: a nonparametric Bayesian approach. *Applied Statistics* 54, 879-895.

McCabe C. Stevens K. Roberts J. Brazier JE. 2005a. Health state values for the HUI2 descriptive system: results from a UK Survey. *Health Economics* 14;231-244

McCabe C. Stevens K. Brazier JE. 2005b. Utility scores for the Health Utilities Index Mark 2: an empirical assessment of alternative mapping functions. *Medical Care* 43;627-635

Sintonen H., 1994. The 15D-measure of health-related quality of life. I. Reliability, validity and sensitivity of its health state descriptive system. National Centre for Health Program Evaluation, Working Paper 41, Melbourne.

Sintonen H., 1995. The 15D-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.

Torrance, G. W., Feeny, D. H., Furlong, W. J. Barr, R. D., Zhang, Y., Wang, Q. A. 1996. Multi-attribute utility function for a comprehensive health status classification system: Health Utilities Index Mark 2. *Medical Care* 34(7); 702-722

Figure 1. Sample mean and predicted health states valuations for the parametric model.

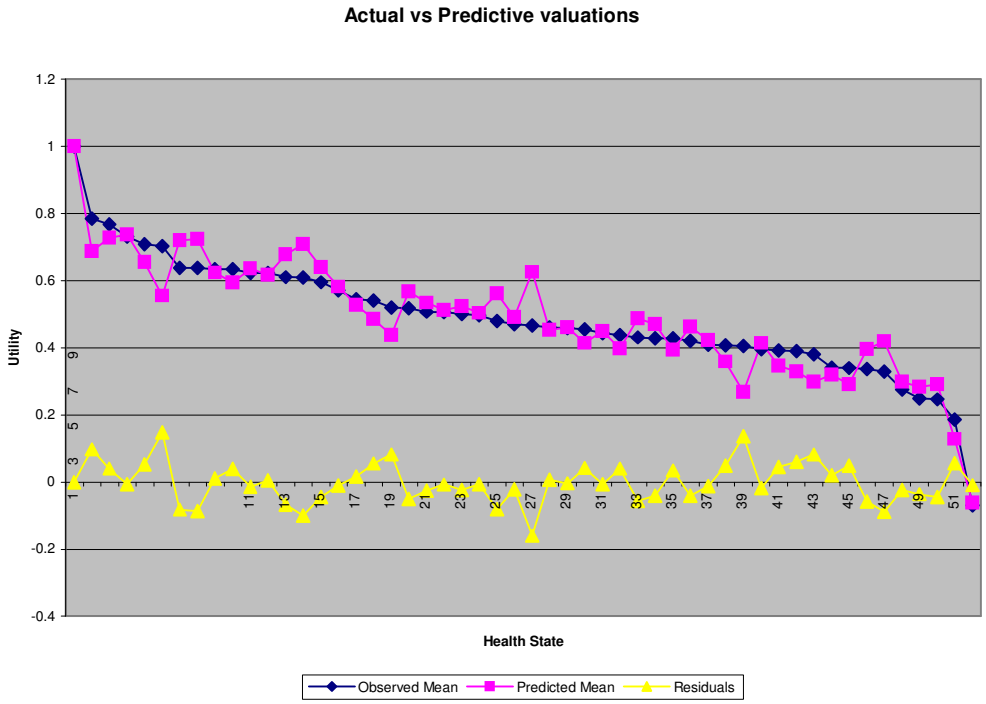

Figure 2. Sample mean and predicted health states valuations for the nonparametric model.

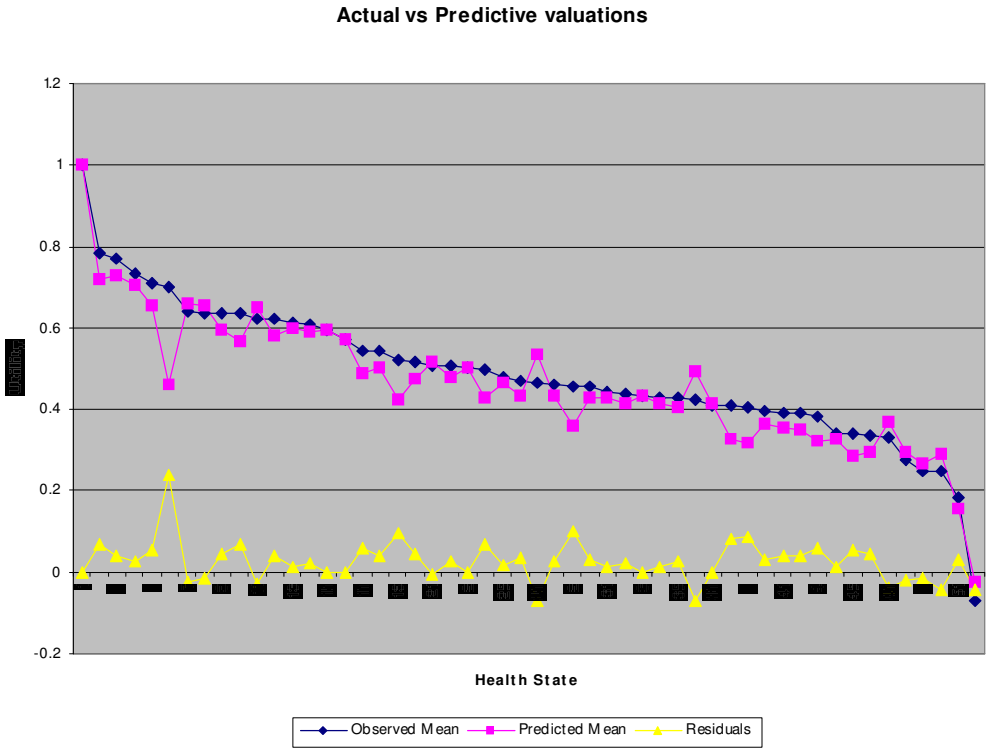

Figure 3. Residuals for the parametric model for each of the 1370 individual health state valuations.

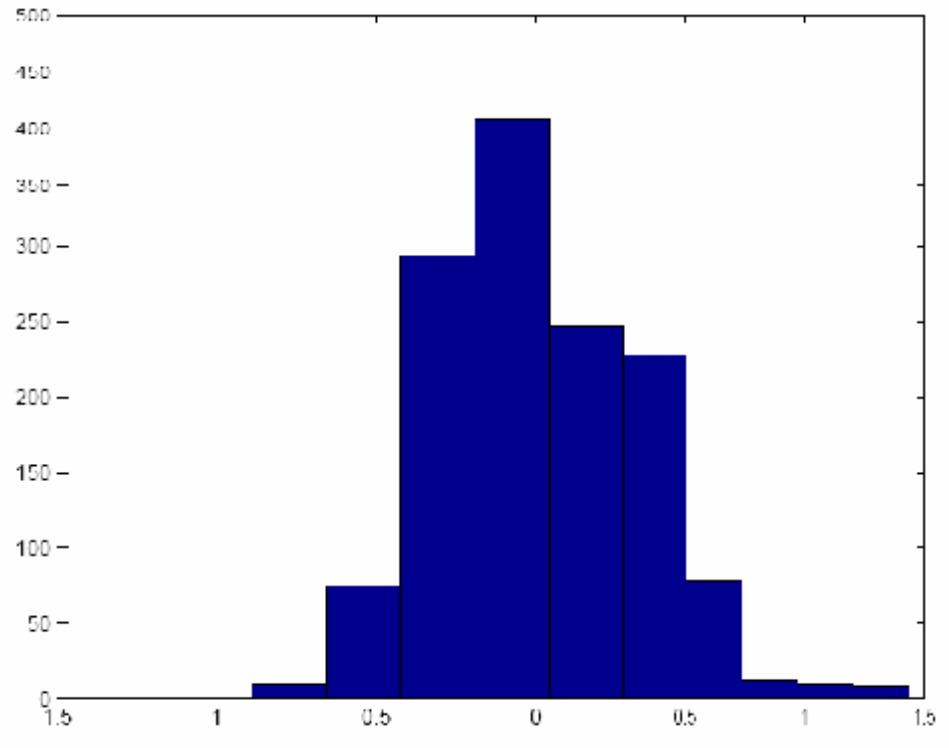

Figure 4. Residuals for the nonparametric model for each of the 1370 individual health state valuation.

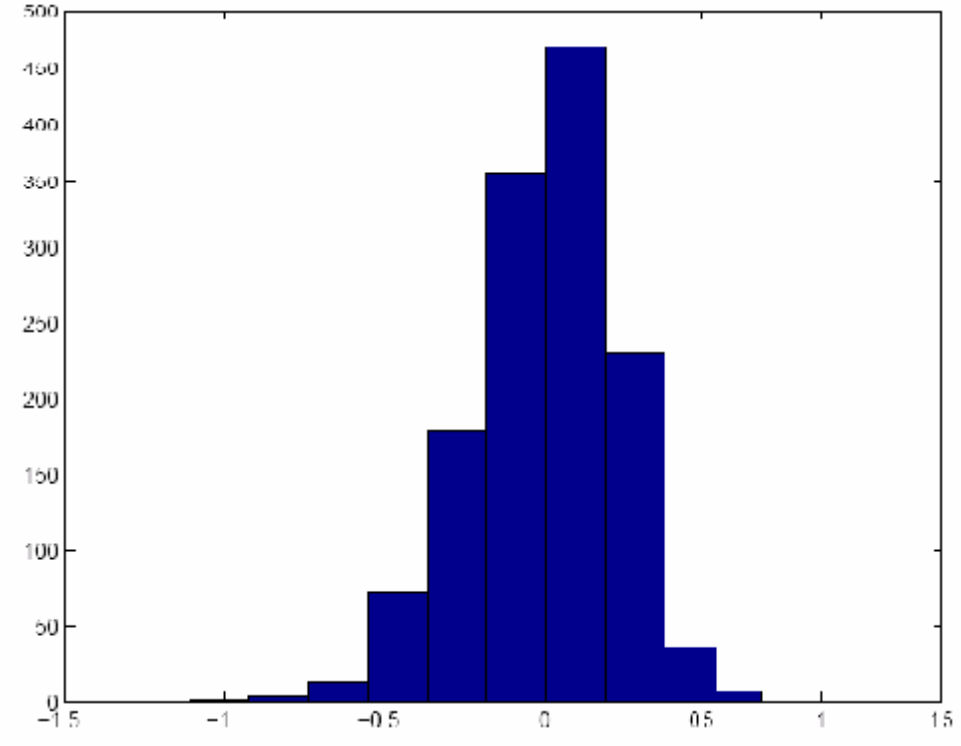

Figure 5. Q-Q plot of standardised predictive errors for the parametric model for the 10 out of sample health states.

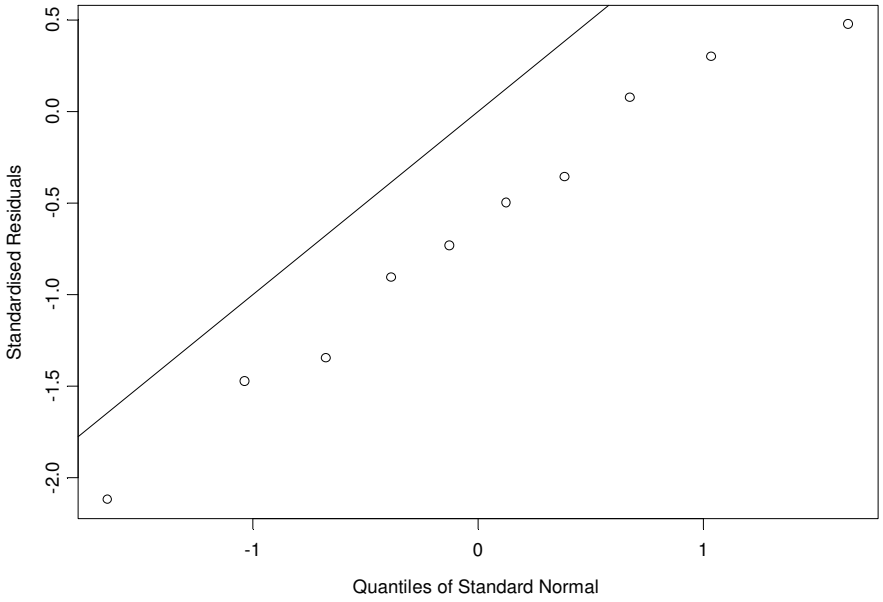

Figure 6. Q-Q plot of standardised predictive errors for the nonparametric model for the 10 out of sample health states.

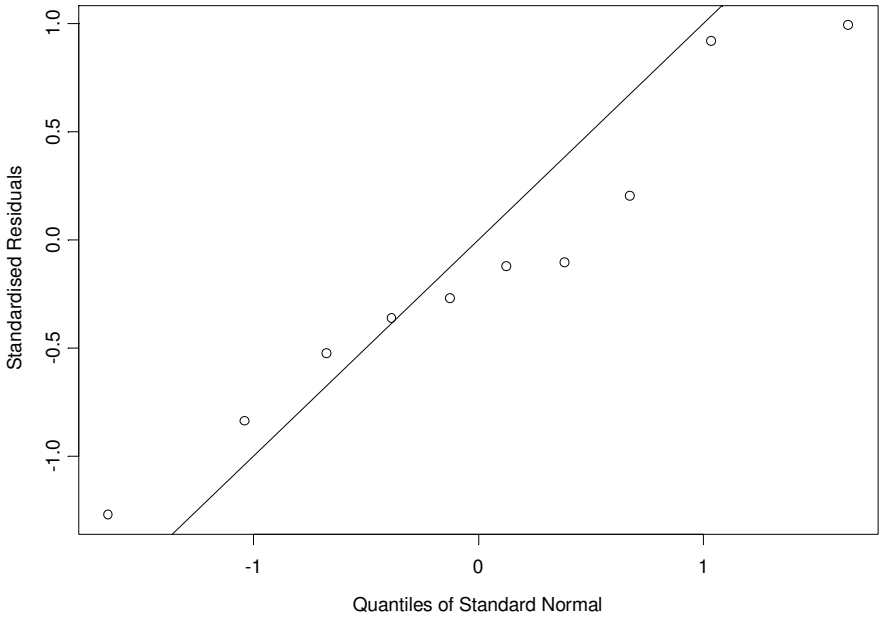

Table 1: The Health Utilities Index Mark 2

Dimension &Levels	Description	Dimension &Levels	Description	Dimension &Levels	Description
Sensation Level 1	Able to see, hear and speak normally for age	**Self Care** Level 1	Eats, bathes, dresses and uses the toilet normally for age	**Emotion** Level 1	Generally happy and free from worry
Level 2	Requires equipment to see or hear or speak	Level 2	Eats, bathes, dresses or uses the toilet independently with difficulty	Level 2	Occasionally fretful, angry, irritable, anxious depressed or suffering from "night terrors"
Level 3	Sees, hears, or speaks with limitations even with equipment	Level 3	Requires mechanical equipment to eat, bathe, dress, or use the toilet independently	Level 3	Often fretful, angry, irritable, anxious depressed or suffering from "night terrors"
Level 4	Blind, deaf, or mute	Level 4	Requires the help of another person to eat, bathe, dress or use the toilet	Level 4	Almost always fretful, angry, irritable, anxious, depressed
				Level 5	Extremely fretful, angry, irritable, anxious or depressed usually requiring hospitalisation usually requiring hospitalisation or psychiatric institutional care
Mobility Level 1	Able to walk, bend, lift, jump and run normally for age	**Cognition** Level 1	Learns and remembers schoolwork normally for age		
Level 2	Walks, bends, lifts, jumps or runs with difficulty but does not require help	Level 2	Learns and remembers schoolwork more slowly than classmates as judged by parents and/or teachers	**Pain** Level 1	Free of pain and discomfort
Level 3	Requires mechanical equipment (such as canes, crutches, braces or a wheelchair) to walk or get around independently	Level 3	Learns and remembers very slowly and usually requires special educational assistance	Level 2	Occasional pain. Discomfort relieved by non-prescription drugs or self-control activity without disruption of normal activities
Level 4	Requires the help of another person to walk or get around and requires mechanical equipment	Level 4	Unable to learn and remember	Level 3	Frequent pain. Discomfort relieved by oral medicines with occasional disruption of normal activities
Level 5	Unable to control or use arms or legs			Level 4	Frequent pain. Frequent disruption of normal activities. Discomfort requires

					prescription narcotics for relief
				Level 5	Severe pain. Pain not relieved by drugs and constantly disrupts normal activities.

Table 2: Inference for the 63 health states

Health State	N	Observed Mean	Non Parametric Posterior inference (No covariates)		Parametric Posterior inference		Non Parametric Posterior inference (Covariates)	
			Mean	S.D.	Mean	S.D.	Mean	S.D.
111232	29	0.6388	0.6616	0.0349	0.7201	0.0427	0.669	0.0339
112123	26	0.7317	0.7068	0.0367	0.7381	0.0473	0.713	0.0362
112222	21	0.7857	0.7174	0.0333	0.6880	0.0495	0.7234	0.0317
121132	25	0.638	0.6533	0.0372	0.7241	0.0427	0.6611	0.0354
122222	11	0.6227	0.6517	0.0339	0.6367	0.0553	0.6594	0.0331
124134	26	0.4702	0.4353	0.0435	0.4918	0.0551	0.4499	0.0417
125211	29	0.6103	0.59	0.0398	0.7096	0.0468	0.5985	0.0382
125425	19	0.3395	0.2852	0.0528	0.2907	0.0650	0.3017	0.0498
132332	23	0.4217	0.4927	0.0391	0.4625	0.0561	0.505	0.0383
133213	22	0.6341	0.5676	0.04	0.5950	0.0576	0.5775	0.0406
133444	28	0.3411	0.3274	0.0449	0.3204	0.0666	0.3444	0.0422
141341	28	0.7089	0.6531	0.042	0.6558	0.0489	0.6622	0.0401
142311	25	0.611	0.599	0.0412	0.6780	0.0449	0.6091	0.0403
144325	22	0.2489	0.2655	0.048	0.2844	0.0642	0.2823	0.0461
212314	28	0.508	0.5162	0.0394	0.5342	0.0514	0.5261	0.0375
213321	23	0.4663	0.5356	0.0395	0.6263	0.0529	0.5459	0.038
214242	25	0.41	0.4127	0.0424	0.4223	0.0613	0.4272	0.0409
221214	21	0.6226	0.5801	0.0408	0.6181	0.0518	0.5883	0.0393
221221	26	0.7683	0.7292	0.0348	0.7275	0.0499	0.7339	0.0336
222142	28	0.5063	0.4814	0.0391	0.5133	0.0582	0.4933	0.0366
222333	25	0.456	0.4266	0.0386	0.4142	0.0633	0.4411	0.0378
223235	22	0.3807	0.3218	0.0435	0.2986	0.0670	0.3389	0.0427
231412	21	0.5417	0.5008	0.0412	0.4858	0.0537	0.5113	0.0391
234111	28	0.5723	0.5725	0.0365	0.5822	0.0485	0.5813	0.0359
235121	22	0.5443	0.4863	0.0421	0.5280	0.0550	0.499	0.0408
241423	23	0.3304	0.3681	0.0446	0.4196	0.0634	0.3822	0.043
242135	26	0.2462	0.2908	0.0449	0.2911	0.0614	0.3082	0.0431
255332	21	0.406	0.3176	0.0498	0.2689	0.0784	0.3353	0.0483
311324	25	0.518	0.4725	0.0411	0.5681	0.0518	0.4848	0.0395
313331	9	0.7028	0.4622	0.0411	0.5552	0.0524	0.4742	0.0399
314431	22	0.3966	0.3646	0.0443	0.4141	0.0555	0.3807	0.0435
315143	22	0.458	0.3596	0.0478	0.4613	0.0596	0.3771	0.0472
315312	23	0.4315	0.4311	0.0439	0.4872	0.0525	0.4432	0.0431
321345	28	0.408	0.3266	0.0456	0.3585	0.0619	0.3411	0.0437
322221	29	0.6353	0.5927	0.0341	0.6244	0.0554	0.6009	0.0339
322412	27	0.4611	0.4347	0.0399	0.4538	0.0574	0.448	0.0385
323331	21	0.4976	0.4285	0.0394	0.5039	0.0579	0.4413	0.0378
331131	23	0.5957	0.5967	0.0389	0.6407	0.0446	0.6056	0.0384
331233	25	0.429	0.4164	0.04	0.4698	0.0599	0.43	0.0381
332225	27	0.275	0.2947	0.0402	0.2986	0.0656	0.3106	0.0391
342124	28	0.5205	0.4239	0.0413	0.4384	0.0596	0.4366	0.0395
343112	27	0.5009	0.5021	0.0386	0.5240	0.0548	0.5135	0.037
344222	20	0.3925	0.3525	0.0435	0.3471	0.0654	0.3663	0.0425
412431	27	0.4389	0.4163	0.0422	0.3983	0.0560	0.4306	0.0409
421114	25	0.481	0.4651	0.0441	0.5621	0.0499	0.4764	0.043

423122	21	0.4429	0.4286	0.0447	0.4494	0.0620	0.4391	0.0424
424313	23	0.3913	0.3507	0.045	0.3302	0.0645	0.3661	0.044
431322	27	0.4287	0.404	0.0407	0.3940	0.0602	0.4177	0.0383
445234	28	0.1857	0.1537	0.048	0.1285	0.0703	0.1718	0.0469
452241	23	0.337	0.2928	0.049	0.3954	0.0771	0.311	0.0485
455445	167	-0.0701	-0.0256	0.0363	-0.0609	0.0878	-0.0018	0.0327
121434*	NA	NA	0.4654	0.0698	0.5056	0.0545	0.4784	0.0676
211223*	NA	NA	0.6424	0.0473	0.6632	0.0537	0.6499	0.0459
224112*	NA	NA	0.5742	0.0518	0.5430	0.0528	0.5824	0.0505
232141*	NA	NA	0.5372	0.0541	0.5525	0.0543	0.5487	0.0522
311124*	NA	NA	0.535	0.062	0.6643	0.0459	0.5449	0.0604
322222*	NA	NA	0.5296	0.0405	0.5141	0.0605	0.5397	0.0398
342223*	NA	NA	0.4015	0.0489	0.4287	0.0656	0.4151	0.0473
341314*	NA	NA	0.3644	0.0677	0.4886	0.0536	0.3773	0.0655
421313*	NA	NA	0.4363	0.0587	0.5116	0.0566	0.4487	0.0571
444335*	NA	NA	0.0993	0.0632	-0.0034	0.0715	0.1187	0.0613
221321*	NA	NA	0.6374	0.0469	0.6867	0.0496	0.6449	0.0455
342411*	NA	NA	0.4305	0.0677	0.4839	0.0542	0.4433	0.0657

* State valued in the validation survey

Table 3: Out of sample predictions for 10 health states

			Nonparametric posterior inference		Parametric inference	
missing state	N	true sample mean	Mean	(s.d.)	mean	(s.d.)
122222	11	0.6227	0.6462	0.0862	0.679	0.1120
133213	22	0.6341	0.5655	0.0747	0.592	0.0874
212314	28	0.5080	0.5157	0.0716	0.613	0.0782
222142	28	0.5063	0.5434	0.0707	0.567	0.0829
235121	22	0.5443	0.5294	0.0742	0.519	0.0865
314431	22	0.3966	0.4247	0.0776	0.390	0.0858
322221	29	0.6353	0.573	0.0629	0.664	0.0791
332225	27	0.2750	0.3647	0.0706	0.354	0.0870
421114	25	0.4810	0.4915	0.0843	0.649	0.0791
452241	23	0.3370	0.4101	0.0873	0.481	0.0978

Figure 7: Conditional posterior distribution functions of the covariates Age and Sex.

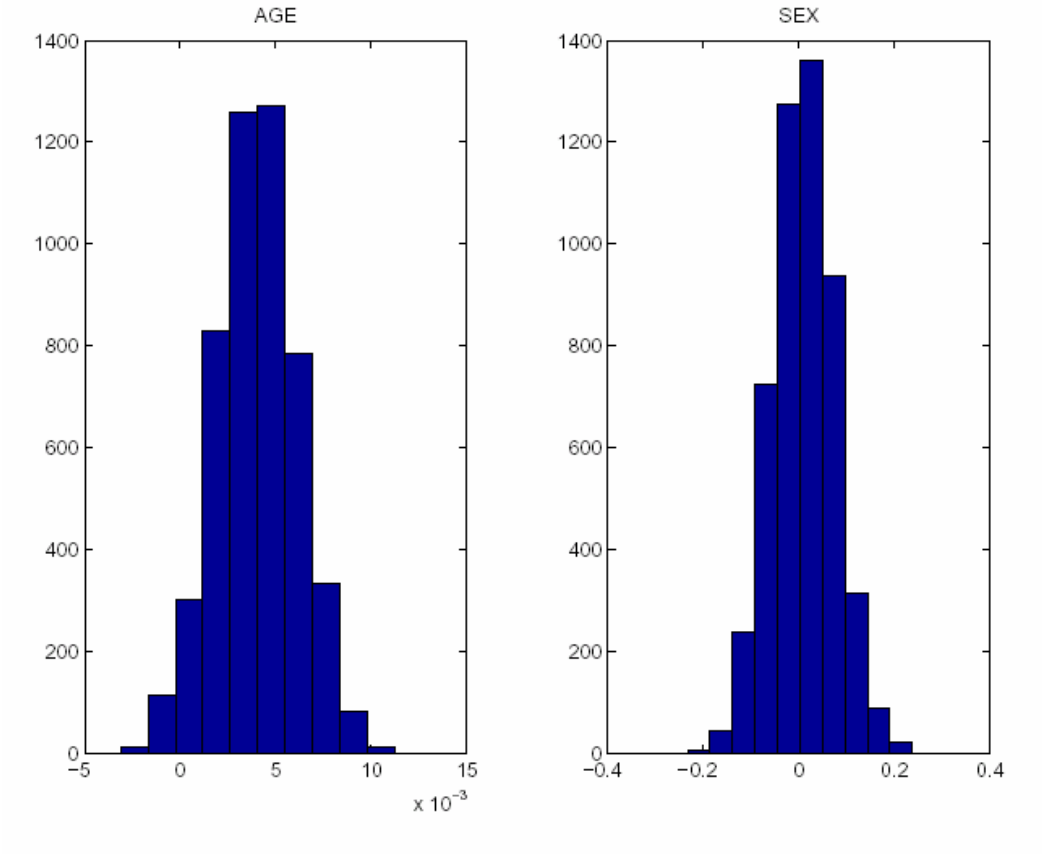

Figure 8. Sample mean and predicted health states valuations for the nonparametric model excluding 313331 health state.

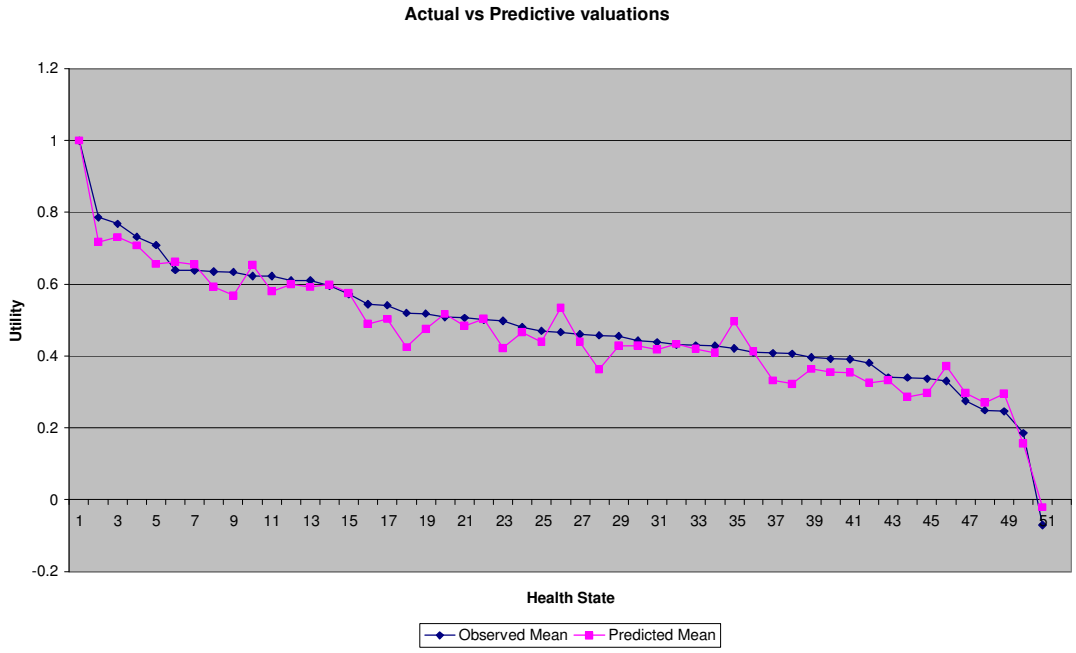

Figure 9. Sample mean and predicted health states valuations for the nonparametric model excluding both 313331 and 122222 health states.