$\mathcal{I}^2\mathcal{SDS}$

The Institute for Integrating Statistics in Decision Sciences

Technical Report TR-2007-11

June 27, 2007

Estimating the Population Utility Function: A Parametric Bayesian Approach

R. Muzaffer Musal
Department of Decision Sciences
The George Washington University, USA

Refik Soyer
Department of Decision Sciences
The George Washington University, USA

Christopher McCabe
Institute of Health Sciences
University of Leeds, UK

Samer A. Kharroubi
Department of Mathematics
University of York, UK

Estimating the Population Utility Function: A Parametric Bayesian Approach

R. Muzaffer Musal
Department of Decision Sciences
The George Washington University, USA

Refik Soyer
Department of Decision Sciences
The George Washington University, USA

Christopher McCabe
Institute of Health Sciences
University of Leeds, UK

Samer A. Kharroubi
Department of Mathematics
University of York, UK

Abstract

Health Related Quality of Life (HRQoL) is a method to measure the perceived physical and mental health over time, as described by the Center for Disease Control. This measure can be used for various purposes such as evaluating the severity of the effect of a disease or comparing treatment methods. There exist alternate HRQoL measurements. In this paper we consider the Health Utility Index Mark II (HUI2) to quantify and describe a population's HRQoL over health states which are composed of multiple attributes. We present a Bayesian framework for population utility estimation and health policy evaluation by introducing a probabilistic interpretation of the multiattribute utility theory (MAUT) approach used in health economics. In so doing, our approach combines ideas from MAUT approach of Raiffa and Keeney (1976) and Bayesian statistics and provides an alternate method of modeling preferences and utility estimation.

Keywords: Bayesian inference; Health economics; Health-related quality of life; Markov chain Monte Carlo methods; HUI2 System.

1. Introduction

Preference based evaluation of health policies have become an integral component of public health policy development in countries such as Australia, Canada and United Kingdom. As pointed out by Brazier (2005), use of preference-based measures of health "… have become a common means of generating health state values for calculating quality-adjusted life years (QALY)."

Use of preference-based measures requires quantification of health state preferences by a sample group of individuals. This preference data is used as a sample to develop an aggregate measure for the population. Authors such as Torrance, Boyle and Horwood (1982) liken this process to determination of a social preference function and present a multi atribute utility theory (MAUT) framework in the sense of Keeney and Kirkwood (1975) and Keeney and Raiffa (1976). The methods that quantify preference based measurement of health (PBMH) are referred to as the health related quality of life measures (HRQoL). These measures are used to quantify a population's preferences over health states as well as of a treatment's effect. Brazier (2005) provided a short list of several common PBMHs and discussed how they are quantified with HRQoL.

Most of the research in this area focuses on development of PBMHs such as Health Utilities Index (HUI) of Torrance, Furlong, Feeny and Boyle (1995), Quality of Well-Being (QWB) scale discussed in Kaplan, Ganiats, Sieber and Anderson (1988) and Short Form (SF-6D) survey discussed in Brazier, Roberts and Deverill (2002). These measures are based on a multi attribute model for evaluating health states using preference weights and scores. They provide a single index number for each health state. Typically an index value "1" denotes perfect health and "0" denotes death. In health economics literature, these index values are referred to as utility. Furthermore, health states have multiple dimensions allowing a multiattribute model. Thus, elicitation of utility requires sophisticated procedures based on standard gambles as reviewed in Brazier and Deverill (1999). A simpler approach for obtaining preference measures is to ask respondents to assign values to health states and convert these to utilities.

Preference based measures such as QWB and SF-6D use what is referred to as the *composite* approach for estimation of the multi attribute utility function for the health states. The composite approach involves direct elicitation of utility of multidimensional health states and as pointed out by Brazier (2005) the approach requires lot more health states than that can be evaluated by a single respondent. Thus, regression models are used with this approach to extrapolate the values of health states, that are not included in the survey.

The alternative method for estimation of the utility function for the health states is the *decomposed* approach employed by the HUI. The decomposed approach uses the MAUT framework developed by Keeney and Raiffa (1976) and determines a functional form for the multi attribute utility function of health states. Based on simplifying assumptions such as *preferential independence* and *utility independence*, the approach yields simpler forms of utility functions and substantially reduces the valuation effort by decomposing the problem into single dimension elicitation problems. Hazen (2004) investigated various independence concepts including those involving the concept of time. The author goes on to describe how the additive or multiplicative decomposition within QALYs can be constructed using these independence concepts and furthermore, discusses how they relate to HUI. In addition to allowing for evaluations for all possible health states, the decomposed approach also provides flexibility in modeling interactions using multiplicative utility functions of Keeney (1974). This is unlike the composite approach where there is no standard method for determining the states required to estimate a model with interaction terms; see Brazier (2005).

It is important to note that both the composite and decomposed approaches provide us with a sample of health state valuation data, that is, with health state utilities from a sample of individuals. The objective is to estimate the health state utilities of the population based on this sample and use the estimated population utility function to evaluate different health policy alternatives. In recognition of this, statistical methods have been considered by earlier researchers such as Dolan (1997) and Brazier et al. (2002). In general these approaches employed linear models with normally distributed error terms. Brazier et al. (2002) went on to include random effect terms. As pointed out by Brazier (2005), these models, that used data from the composite approach, "have estimated crude summary terms for interactions" and have required range of transformations to deal with highly skewed data.

More recently a nonparametric Bayesian approache has been considered in Kharroubi, Brazier, O'Hagan and Roberts (2007) and Kharroubi, O'Hagan and Brazier (2005) for estimation of the HRQoL of a population. Utility quantifies the HRQoL measured by the SF-6D measurement which is based on six attributes, namely, physical functioning, role limitations, social functioning, pain, mental health and vitality where each attribute has four to six levels. The specification of the six attributes levels describe a health state where 18,000 such possible health states are possible. The authors treat health states as single attribute objects and utilize standard gamble techniques to elicit utility from individuals. Since it would be infeasible to construct a utility function over all health states using the standard gamble technique, the authors propose a sampling method where an individual goes through the standard gambles for a limited number of health states. Utility function is estimated from these health states using a multivariate nonparametric Bayesian model.

Even though this model has desirable properties, it is based on the composite approach that requires large number of evaluations of health states. As noted by Brazier (2005), although the composite approach involves cognitively complex tasks in elicitation, most of the literature on statistical modeling of health state data is limited to this approach. In this paper we will consider parametric Bayesian models and inference to analyze health state utility data using a decomposed approach. In so doing, we describe uncertainty about the unknown utility function probabilistically and pose the population utility estimation problem as a Bayesian inference problem. Our framework is different than Kharroubi, O'Hagan and Brazier (2005) due to its parametric nature and its use of a decomposed approach based on a multiattribute utility function.

A synopsis of our paper is as follows. In Section 2 we present a model for describing uncertainty about single attribute utility functions and discuss Bayesian estimation of the single attibute population utility function. Using a decomposed approach based a multiplicative multiattribute utility function Section 3 discusses estimation of population multiattribute utility functions and considers probabilistic evaluation of health state preferences. Extension of the utility models are discussed in Section 4 where covariate and heterogeneity effects are incorporated into the analysis. Implementation of the Bayesian framework is presented in Section 5 using actual health state preference data. Concluding remarks are given in Section 6.

2. Bayesian Modeling of Utility Functions

Torrance, Boyle and Harwood (1982) presented a MAUT approach for quantification of society's preferences for health states. The authors used a multiattribute utility model to aggregate sample individuals' utility and estimated society's preferences over health states and provided a framework for the development of a popular HRQoL measure, the Heath Utility Index (HUI). Torrance et al. (1996) outlined the construct of Health Utilities Mark 2 (HUI2) measurement which will be used in our development. HUI2 has seven attributes with 3-5 levels, (sensation, mobility, emotion, cognition, self-care, pain, and fertility) that describes around 24,000 health states. The authors used the multiplicative multiattribute model in order to estimate the utility of these health states. The Bayesian framework, that we present in the sequel, is motivated by the Torrance et al. (1996) setup.

We consider a multiattribute evaluation problem with K+1 alternatives and C attributes where the interest is in estimating the utility of these alternatives for a population based on utilities obtained from a sample of N individuals. In the context of HUI systems discussed above, the K+1 alternatives represent the health states associated with different health policy initiatives and C attributes represent the multiple dimensions considered in the MAU model of the HUI.

Our development requires a probability model for describing population utility. Following the decomposed approach of Torrance et al. (1996), as is common in the HUI literature, we will use multiplicative utility functions. Use of such forms for utility functions can be justified in the presence mutual utility independence of the attributes. This particular form allows us to model population utility for each of the attributes and estimate population utility on a given dimension. Once such estimation is completed for all attributes independently, population's MAU function can be obtained via the multiplicative utility model.

2.1 A Model for Population Utilities

We assume that preference ordering of the $K+1$ alternatives with respect to each attribute is known for the population, that is, we assume that the preference ordering of the population of individuals is identical and monotonic. Let X denote the level of a single attribute c, then it follows that

$$X_1 \prec X_2 \prec X_3 \prec \cdots \prec X_K \prec X_{K+1}.$$

In this setup we are interested in making inference about the unknown population utilities $u(X_2) < \cdots < u(X_K)$, where $u(X_1) = 0$ and $u(X_{K+1}) = 1$. We may have a prior opinion on these values and we are interested in updating this prior opinion based on the sample utility measurements on the N individuals. In what follows, the terms value and utility will be used interchangeably.

We define $u(X_{c,j}) = u_{c,j}$ and let $u^i_{c,j}$ denote the utility declared by the $i-th$ individual for attribute c at level j. For general purposes, we also define $K_c + 1$ to denote the number of levels of attribute c. The attribute label is supressed unless it becomes necessary. We focus on a single criterion and to reflect the ordering $u(X_1) < u(X_2) < \cdots < u(X_K) < u(X_{K+1})$, that applies to the population, we assume that for all individuals

$$0 < u_2 < u_3 < \cdots < u_K < 1. \quad (1)$$

The first step in our development is to consider a probability model for the utility vectors $\boldsymbol{u}^i = (u_1^i,\, u_2^i,\, \ldots, u_K^i)$ which is consistent with the ordering given by (1) and which is flexible enough to reflect the diminishing utility scenario encountered in many applications. One such multivariate model is the ordered Dirichlet distribution which has been used in reliability growth modeling by van Dorpe, Mazzuchi and Soyer (1997) and by Erkanli, Mazzuchi and Soyer (1998). The ordered Dirichlet model for $\boldsymbol{u} = (u_2, u_3,$ $\ldots, u_K)$ is given by

$$p(\boldsymbol{u} \mid \beta, \boldsymbol{\alpha}) = \frac{\Gamma(\beta)}{\prod_{\mathrm{j}=2}^{K+1} \Gamma(\beta\alpha_{\mathrm{j}})} \prod_{\mathrm{j}=2}^{K+1} (u_j - u_{j-1})^{\beta\alpha_{\mathrm{j}}-1}, \tag{2}$$

where $u_1 = 0$, $u_{K+1} = 1$, β and $\boldsymbol{\alpha} = (\alpha_2, \alpha_3,\, \ldots, \alpha_{K+1})$ are the parameters of the model such that $\alpha_j > 0$ and $\sum_{j=2}^{K+1} \alpha_j = 1$. Note that the distribution is defined over the simplex $\left\{\boldsymbol{u} \mid 0 \leq u_2 \leq\, u_3 \leq \cdots \leq u_K \leq 1\right\}$ which is consistent with the restrictions in (1). The model implies that the changes in the utilities $(u_j - u_{j-1})$, for $j = 2, \ldots, K+1$, follow a Dirichlet distribution.

As shown by van Dorpe, Mazzuchi and Soyer (1997), the marginal distributions of the ordered Dirichlet model are beta densities, that is,

$$p(u_j \mid \beta, \boldsymbol{\alpha}) = \frac{\Gamma(\beta)}{\Gamma(\beta\alpha_j^*)\,\Gamma[\beta(1-\alpha_j^*)]}\; (u_j)^{\beta\alpha_j^*-1}\,(1-u_j)^{\beta(1-\alpha_j^*)-1} \tag{3}$$

for $j = 2, \ldots K$, where $\alpha_j^* = \sum_{k=2}^{j} \alpha_k$. We will denote the beta density of (3) as

$$(u_j \mid \beta, \boldsymbol{\alpha}) \sim Beta\Big(\beta\alpha_j, \beta(1-\alpha_j^*)\Big). \tag{4}$$

Note that the above implies that

$$E[u_j \mid \beta, \boldsymbol{\alpha}] = \alpha_j^* \text{ and } V[u_j \mid \beta, \boldsymbol{\alpha}] = \frac{\alpha_j^*(1-\alpha_j^*)}{\beta+1} \tag{5}$$

where β is a *degree of belief* parameter with lower values of β reflecting more spread in the distribution. Furthermore, it can be shown that [see van Dorpe et al. (1997)]

$$(u_j - u_i) \mid \beta, \boldsymbol{\alpha} \sim Beta\Big(\beta(\alpha_j^* - \alpha_i^*), \beta(1-\alpha_j^*+\alpha_i^*)\Big) \text{ for } i < j, \tag{6}$$

and thus, we obtain

$$E[u_j - u_{j-1} \mid \beta, \boldsymbol{\alpha}] = (\alpha_j^* - \alpha_{j-1}^*) = \alpha_j. \tag{7}$$

The above distributions provide meaningful interpretations for the parameters α_j and α_j^* in terms of properties of the utility function. Specifically, α_j can be interpreted as the expected increase in utility as a result of going from attribute level X_{j-1} to attribute level X_j and and α_j^* is the expected utility at attribute level X_j. Furthermore, it is easy to see that α_j^* is increasing with j, implying that for the population we expect utility is an increasing function of the attribute when high values of the attribute are desirable. Also, if $E[u_j - u_{j-1} \mid \beta,\ \boldsymbol{\alpha}] = \alpha_j$ is a decreasing sequence in j, then we expect that the marginal utility is diminishing as the attribute level gets larger. This is equivalent to requiring that $E[u_j \mid \beta,\ \boldsymbol{\alpha}] = \alpha_j^*$ is discrete concave in j, that is, we expect that individuals in the population is risk averse with respect to the particular attribute. By choosing α_j's differently we can represent different attitudes towards risk. For example, by specifying α_j as an increasing sequence in j, we represent a risk seeking behavior. Similarly if $\alpha_j = \alpha$ for all j, then we have the risk neutral behavior. Note that we may have prior beliefs about the general behavior of the expected utility function and we can incorporate that in our Bayesian analysis. Such prior beliefs can be used in specification of the prior distribution of the parameters $\boldsymbol{\alpha}$ and β. Thus, the ordered Dirichlet distribution (2) provides a flexible model to describe uncertainty about utility.

2.2 Bayesian Estimation of Population Utilities

In the Bayesian setup we are interested in describing uncertainty about the population utility function $\boldsymbol{u} = (u_2,\ \ldots, u_K)$ based on the the information provided by the sample of N utility vectors $\boldsymbol{u}^i = (u_2^i,\ u_3^i,\ \ldots, u_K^i),\ i = 1, \ldots, N$ from the ordered Dirichlet distribution in (2). Given sample utilities given $\boldsymbol{u}^1, \boldsymbol{u}^2, \ldots, \boldsymbol{u}^N$ from the N individuals, we update our knowledge about $\boldsymbol{u}$ via the calculus of probability. In the Bayesian framework this is achieved by obtaining the posterior predictive distribution $p(\boldsymbol{u}^{N+1}|\boldsymbol{u}^1, \boldsymbol{u}^2, \ldots, \boldsymbol{u}^N)$. Thus, the objective of this treatment is to obtain the posterior distribution for the population's utility function. In what follows, we will pose this as an inference problem and present the Bayesian machinery to obtain $p(\boldsymbol{u}^{N+1}|\boldsymbol{u}^1, \boldsymbol{u}^2 ...,\ \boldsymbol{u}^N)$.

In order to obtain the posterior predictive distribution of $\boldsymbol{u}$, we need to specify our prior distribution of the unknown population parameters $(\beta, \boldsymbol{\alpha})$. We denote this distribution by $p(\beta,\ \boldsymbol{\alpha}|\mathcal{H})$ where $\mathcal{H}$ represents prior beliefs about the utility function of the population. Note that in specifying the prior distribution, we can use the results (5)-(7) as well as the implications of monotonocity of α_j's. If such a monotonocity assumption is used then the joint distribution of α_j's should reflect that. For example, in such case an ordered Dirichlet distribution can be used as the joint distribution of α_j's. In our setup, we assume that β and $\boldsymbol{\alpha}$ are conditionally independent given $\mathcal{H}$. We use a Dirichlet distribution to model the $\boldsymbol{\alpha}$ parameters and specify a gamma density for β. Once the prior is specified, we can revise our uncertainty about unknown parameters (β, $\boldsymbol{\alpha}$) via the Bayes rule, that is,

$$p(\beta, \boldsymbol{\alpha}|\boldsymbol{u^N}) = \frac{L(\beta, \boldsymbol{\alpha}; \boldsymbol{u^N})\, p(\beta, \boldsymbol{\alpha}|\mathcal{H})}{\int_{\beta,\boldsymbol{\alpha}} L(\beta, \boldsymbol{\alpha}; \boldsymbol{u^N})\, p(\beta, \boldsymbol{\alpha}|\mathcal{H})\, d\boldsymbol{\alpha}\, d\beta} \tag{8}$$

where $L(\beta, \boldsymbol{\alpha}; \boldsymbol{u^N})$ is the likelihood function of $(\beta, \boldsymbol{\alpha})$ given $\boldsymbol{u^N} = (\boldsymbol{u}^1, \boldsymbol{u}^2, \ldots, \boldsymbol{u}^N)$.

Since the utilities $(\boldsymbol{u}^1, \boldsymbol{u}^2, \ldots, \boldsymbol{u}^N)$ are conditionally independent vectors from the the ordered Dirichlet distribution (1). Given N, K dimensional utility vectors, we can write the likelihood function of the parameters $(\beta, \boldsymbol{\alpha})$ as

$$L(\beta, \boldsymbol{\alpha}; \boldsymbol{u^N}) = \prod_{i=1}^{N} \left[\frac{\Gamma(\beta)}{\prod_{j=2}^{K+1} \Gamma(\beta\alpha_j)} \prod_{j=2}^{K+1} (u_j^i - u_{j-1}^i)^{\beta\alpha_j - 1} \right] \tag{9}$$

by using the ordered Dirichlet model.

The posterior density $p(\beta, \boldsymbol{\alpha}|\boldsymbol{u^N})$ in (8) can not be obtained analytically and thus requires use of iterative simulation methods known as Markov chain Monte Carlo (MCMC) techniques. Note that evaluation of (8) involves the Bayesian analysis of the ordered Dirichlet model given by (2). Previous uses of the ordered Dirichlet model in the literature were generally limited to prior distributions, but in this case the ordered Dirichlet model is used as the sampling model (or the likelihood) for the utility vectors.

Once the posterior density $p(\beta, \boldsymbol{\alpha}|\boldsymbol{u^N})$ is evaluated, the next step is to obtain the posterior distribution of the utility function for the population, that is, we need to obtain the joint distribution $p(\boldsymbol{u}|\boldsymbol{u}^1, \boldsymbol{u}^2, \ldots, \boldsymbol{u}^N) = p(\boldsymbol{u}|\boldsymbol{u^N})$. In other words, we need to evaluate the posterior predictive distribution

$$\boldsymbol{p}(\boldsymbol{u}|\boldsymbol{u^N}) = \int_{\beta,\,\boldsymbol{\alpha}} p(\boldsymbol{u}|\beta, \boldsymbol{\alpha}, \boldsymbol{u^N})\, p(\beta, \boldsymbol{\alpha}|\boldsymbol{u^N})\, d\boldsymbol{\alpha}\, d\beta, \tag{10}$$

which reduces to

$$p(\boldsymbol{u}|\boldsymbol{u}^N) = \int_{\beta,\,\boldsymbol{\alpha}} p(\boldsymbol{u}|\beta,\boldsymbol{\alpha})\,p(\beta,\boldsymbol{\alpha}|\boldsymbol{u}^{\boldsymbol{N}})\,d\boldsymbol{\alpha}\,d\beta, \tag{11}$$

due to the independence of $\boldsymbol{u}$ and $\boldsymbol{u}^N$ given $(\beta,\boldsymbol{\alpha})$. Note that in the above $p(\boldsymbol{u}|\beta,\boldsymbol{\alpha})$ is the ordered Dirichlet density shown in (2). Again the posterior predictive distribution can not be obtained analytically but given the posterior samples from $p(\beta,\boldsymbol{\alpha}|\boldsymbol{u_N})$, (11) can be approximated by the Monte Carlo integral

$$p(\boldsymbol{u}|\boldsymbol{u}^N) \approx \frac{1}{S}\sum_{s=1}^{S} p(\boldsymbol{u}|\beta^{(s)},\boldsymbol{\alpha}^{(\boldsymbol{s})}), \tag{12}$$

where $\left(\beta^{(s)},\boldsymbol{\alpha}^{(\boldsymbol{s})}\right)_{s=1}^{S}$ are samples from the posterior density $p(\beta,\boldsymbol{\alpha}|\boldsymbol{u}^{\boldsymbol{N}})$. The posterior predictive distribution $p(\boldsymbol{u}|\boldsymbol{u}^N)$ represents probability distribution of the population's utility function.

3. Estimation of MAU Function and Evaluation of Health States

The above development was presented for a single attribute, that is, for attribute c, with K_c+1 levels and observed utility vectors $\boldsymbol{u}_{\boldsymbol{c}}^i = (u_{c,1}^i, u_{c,2}^i, \ldots, u_{c,K}^i)$ for individuals $i=1,\ldots,N$. Note that in the above development attribute index c was suppressed. In general for specifying the ordered Dirichlet model, all model parameters are supposed to be indexed by c, that is, we have $(\beta_c,\boldsymbol{\alpha_c})$ with prior $p(\beta_c,\boldsymbol{\alpha_c}|\mathcal{H})$ for attribute c. Using the Bayesian machinery our development in Section 2 resulted in the posterior utility distribution $p(\boldsymbol{u_c}|\boldsymbol{u_c^N})$ for attribute c. The proposed model is justifiable when we have mutual utility independence of the attributes [see for example, Keeney and Raiffa (1976)]. Then the above development can be extended to C attributes and the posterior utility distributions $p(\boldsymbol{u_c}|\boldsymbol{u_c^N})$ are obtained for $c=1,\ldots,C$. In this case the parameters $(\beta_c,\ \boldsymbol{\alpha_c})$ are assumed to be independent for $c=1,\ldots,C$ and thus our approach yields independent posteriors $p(\boldsymbol{u_c}|\boldsymbol{u_c^N})$ for $c=1,\ldots,C$.

A common method of decomposition that is used to account for potential interactions in attribute utilities is the multiplicative utility model; see Brazier (2005). The multiplicative model is specified as,

$$1+ku(X_{1,A_i},\ldots,X_{C,A_i}) = \prod_{c=1}^{C}(1\,+\,kk_{\mathrm{c}}u(X_{A_i})), \tag{13}$$

where $0<k_c<1\,and\,k>-1.$

In the above k, and k_c are scaling coefficients that fixes the range of utility between 0 and 1. We assume that, given their respective parameters, the k_c's are independently beta distributed namely,

$$p(k_c \mid \kappa, \nu_c) \sim Beta\Big(\kappa\nu_c, \kappa(1-\nu_c)\Big). \tag{14}$$

In the above ν_c is the expected value of scaling coefficient k_c, that is, $0 < \nu_c < 1$ for $c = 1, \ldots, C$. Parameter κ represents the analyst's strength of belief about specification of ν_c. By specifying independent priors for ν_c but assuming a common belief parameter κ for all attributes in (14) we can describe dependence of k_c's in the multiplicative model. Once the joint prior $p(\boldsymbol{\nu}, \kappa|\mathcal{H})$, where $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_C)$, is specified we can obtain the posterior distribution of $\boldsymbol{\nu}$ and κ using the Bayes' rule. More specifically

$$p(\boldsymbol{\nu}, \kappa|\boldsymbol{K^N}) \propto L(\kappa, \boldsymbol{\nu}; \boldsymbol{K^N})p(\boldsymbol{\nu}, \kappa|\mathcal{H}),$$

where $L(\kappa, \boldsymbol{\nu}; \boldsymbol{K^N})$ is the likelihood function of $\boldsymbol{\nu}$ and κ given the sample of scaling coefficients $\boldsymbol{K^N} = \big(\boldsymbol{k^i} = (k_1^i, \ldots k_C^i); i = 1, \ldots, N\big)$ elicited from N individuals. The likelihood function can be written as the product of beta densities, that is,

$$L\big(\kappa, \boldsymbol{\nu}; \boldsymbol{K^N}\big) = \prod_{i=1}^{N}\prod_{c=1}^{C} \frac{\Gamma(\kappa)}{\Gamma(\kappa\nu_c)\Gamma(\kappa(1-\nu_c))}\big(k_c^i\big)^{\kappa\nu_c - 1}\big(1 - \big(k_c^i\big)\big)^{\kappa(1-\nu_c)-1} \tag{15}$$

For any form of the prior distribution on $(\kappa, \boldsymbol{\nu})$, the posterior distribution $p(\kappa, \boldsymbol{\nu}|\boldsymbol{K^N})$ can not be obtained analytically. As in the other cases we will develop the Bayesian analysis using MCMC methods.

Once $p(\kappa, \boldsymbol{\nu}|\boldsymbol{K^N})$ is obtained, we can evaluate the posterior predictive distribution of the population vector $\boldsymbol{K}$ via

$$p\big(\boldsymbol{K} \,|\, \boldsymbol{K^N}\big) = \int_{\kappa, \boldsymbol{\nu}} p(\boldsymbol{K}|\kappa, \boldsymbol{\nu})\, p(\kappa, \boldsymbol{\nu}|\boldsymbol{K^N})\, d\kappa\, d\boldsymbol{\nu}, \tag{16}$$

The above can be approximated by the Monte Carlo integral

$$p(\boldsymbol{K} \,|\, \boldsymbol{K^N}) \approx \frac{1}{S}\sum_{s=1}^{S} p(\boldsymbol{K} \,|\, (\kappa, \boldsymbol{\nu})^{(\boldsymbol{s})}) \tag{17}$$

using S realizations from the posterior distribution $p(\kappa, \boldsymbol{\nu}|\boldsymbol{K^N})$. For parameter k, we have the equality

$$1 + k = \prod_{c=1}^{C}(1 + kk_c),$$

Thus, given the realizations of $\boldsymbol{K}^{(s)}$, $s = 1, \ldots, S$, for each realization the above identity can be solved for the interaction parameter k and its posterior predictive distribution can be obtained. We note that if the posterior predictive distribution of k is concentrated around 0 then this is an indication of support for an additive utility model for the population.

In the above we have presented how to obtain two sets of posterior samples, that is, sample of utilities and weights. Thus, we can make probability statements using the multiattribute utility function in (13). Given a specific health state, say A_i, we can evaluate the corresponding multiattribute function using the posterior samples associated with the utilities corresponding to attribute levels. Based on these for a given health state A_i we can obtain the posterior distribution of $u(X_{1,A_i}, X_{2,A_i}, \ldots, X_{\boldsymbol{C},A_i})$ using a Monte Carlo estimate. Thus, we can make probability statements on whether health state A_i is preferred to A_j, that is, $Pr\{A_i \succ A_j | \boldsymbol{K^N}, \boldsymbol{u_1^N}, \ldots, \boldsymbol{u_C^N}\}$. This probability is equivalent to

$$Pr\Big\{u\left(X_{1,A_i}, X_{2,A_i}, \ldots, X_{\boldsymbol{C},A_i}\right) > u\Big(X_{1,A_j}, X_{2,A_j}, \ldots, X_{\boldsymbol{C},A_j}\Big) | \boldsymbol{K^N}, \boldsymbol{u_1^N}, \ldots, \boldsymbol{u_C^N}\Big\} \tag{18}$$

which can be approximated using the MCMC samples. In other words, it is possible to compute the posterior probability that a particular health state is preferred to another health state in the population.

4. Incorporating Covariate Effects and Heterogeneity

Our development above can be modified to incorporate the effects of covariates on utility at a given attribute level. In so doing, we model the expected changes in utility between the adjacent levels of an attribute. In what follows, we suppress the attribute subscript and use a logit transform for individual i's expected utility change from attribute level $j-1$ to j as

$$\eta_j^i = log\left(\frac{\alpha_j^i}{\alpha_{K+1}^i}\right), \; j = 2, \ldots, K. \tag{19}$$

Note that the expected utility change from level K to $K+1$ is used as the level of reference in (19). The above relationship is defined for each attribute $c = 1, \ldots, C$ separately. For attribute c we can write

$$\begin{pmatrix} \eta_2^i \\ . \\ . \\ . \\ \eta_K^i \end{pmatrix} = \begin{pmatrix} \chi_2 \\ . \\ . \\ . \\ \chi_K \end{pmatrix} + z_1^i \begin{pmatrix} \rho_{2,1} \\ . \\ . \\ . \\ \rho_{K,1} \end{pmatrix} + ... + z_Q^i \begin{pmatrix} \rho_{2,Q} \\ . \\ . \\ . \\ \rho_{K,Q} \end{pmatrix} \tag{20}$$

for Q covariates $z_1, \ldots, z_Q$.

Note that in analyzing the above model, the likelihood function of the unknown parameters will be based on the ordered Dirichlet model (2). Typically, the prior distributions for parameters $\boldsymbol{\chi} = (\chi_2, \ldots \chi_K)$ and $\boldsymbol{\rho} = (\rho_{ij}; i = 2, \ldots, K; j = 1, \ldots, Q)$ will be specified as independent multivariate normal distributions. As before for any choice of prior distributions, evaluation of the posterior distributions requires use of MCMC methods. Note that the model can be generalized to C attributes.

The above model can be easily extended to capture heterogeneity in utility change by introducing random effect terms associated with each individual. More specifically we can write (20) as

$$\begin{pmatrix} \eta_2^i \\ . \\ . \\ . \\ \eta_K^i \end{pmatrix} = \begin{pmatrix} \chi_2 \\ . \\ . \\ . \\ \chi_K \end{pmatrix} + z_1^i \begin{pmatrix} \rho_{2,1} \\ . \\ . \\ . \\ \rho_{K,1} \end{pmatrix} + ... + z_Q^i \begin{pmatrix} \rho_{2,Q} \\ . \\ . \\ . \\ \rho_{K,Q} \end{pmatrix} + \begin{pmatrix} \epsilon_2^i \\ . \\ . \\ . \\ \epsilon_K^i \end{pmatrix}, \tag{21}$$

where random component $\boldsymbol{\epsilon_i} = (\epsilon_2^i \ldots \epsilon_K^i)$ is assumed to have the multivariate normal distribution as

$$\boldsymbol{\epsilon_i} \sim MVN(\mathbf{0}, \boldsymbol{W_\epsilon}),$$

with unknown covariance matrix $\boldsymbol{W_\epsilon}$. Given $\boldsymbol{W_\epsilon}$, $\boldsymbol{\epsilon_i}$'s are conditionally independent random quantities across the individuals.

The above setup can be adopted for the scaling weights k_c's of the multiplicative utility model. More specifically, we can introduce the logit transformation on the expected scaling factors ν_c as

$$\eta_c^i = log\left(\frac{v_c^i}{1 - v_c^i}\right) = \mu_c + \lambda_{1,c} z_1^i + \ldots + \lambda_{Q,c} z_Q^i \tag{22}$$

for $c = 1, \ldots C$. The above can be extended by including an additive random effect term ϵ_c^i to incorporate unknown heterogeneity. As in the previous case the Bayesian analysis requires use of MCMC methods.

In our implementations in the next section, we use BRugs**,** a package within the R software, to apply MCMC methods by running the software WinBUGS.

5. Illustration using HUI2 based Data

In this section we illustrate the implementation of the models introduced in previous sections using data from McCabe, Stevens and Brazier (2004). The data is from an HUI2 based survey conducted on N= 201 individuals drawn from the general population of UK. The purpose of the McCabe et. al. paper is to compare the various methods that are used to convert the data obtained from visual analog scale (VAS) into utility scale in order to "explore the implications for health care resource allocation". The sample was stratified by mainland UK socio-economic region, based on the 1991 census. Marital status, gender and age were three of the characteristics that were monitored.

5.1 Description of the Data

HUI2 is originally developed with seven attributes which describes an individual's health state. One of these attributes, fertility, is not included in this survey. The attributes self care, sensation and cognition has four while mobility, emotion and pain has five levels which is used to describe various situations within these attributes. The levels of the attributes are structured such that for most people there is a natural order where the smaller index level is preferred to its higher adjacent level. Furthermore level 1 of all attributes are fixed at a value of 100 and the highest indexed attribute level is set at 0. We scale these values between 0 and 1. Note that the order of attribute levels is reversed in our development. For instance $(K+1)$st level describes the best attribute value, whereas in the data it represents the worst level.

In order to construct a MAU model, in addition to the utility assigned to each attribute level we need the attribute scaling weights. In conventional MAUT the weights are obtained by eliciting a value for the state where an attribute has the highest and the others the lowest possible values. The same method is not applied in this survey as these would be harder for the respondents to evaluate. Instead they are prompted to provide values for health states where an attribute is set at its worst while the others at their best levels. These are referred to as the *corner health states*. The authors' model is constructed on the disvalue scale (1 – utility) where the worst health state is assigned, a value of 1, and the best, a value of 0. In order to evaluate the utility of a health state the authors simply subtract from 1 the disutility of that health state. The respondents were asked to provide a value between 0 and 100 for the 4 corner health states, namely for the attributes pain, cognition, sensation and emotion. These values are transformed to disutility by subtracting from one. The weights of self care and mobility are obtained using the method described in Torrance et al. (1996) and McCabe et al. (2004). Estimating these attributes weights are problematic since they do not have "structural independence". We apply the same estimation method as well.

In our analysis we only consider values from respondents whose utility increase with smaller indexes of the health attributes. Therefore the number of observations we have in each attribute is not the same. We provide descriptive statistics in Table 1. We eliminate from the original sample, individuals who did not assign a value of 100 to the best health state. Furthermore, when the health states, where both mobility and self care attributes were set at their worst level and where mobility was set at its intermediate level were compared, if an individual provided a lower score for the latter health state (s)he was removed from the study. There were 148 individuals remaining from the original 201. When the two unknowns, the interaction term and the disutility of self care are solved for, 88 individuals are eliminated from whom we could not get valid solutions. This leaves a sample size of 60 for the modelling of attribute weights. In table 2 we report the summary statistics of the disutility weights of the attributes for the multiplicative utility model.

Table 1: Descriptive statistics of attribute levels.

	$Sensation$		$Mobility$			$Emotion$			$Cognition$		$SelfCare$		$Pain$		
	u_2	u_3	u_2	u_3	u_4	u_2	u_3	u_4	u_2	u_3	u_2	u_3	u_2	u_3	u_4
$Mean$	.34	.63	.26	.48	.74	.19	.39	.64	.34	.62	.29	.61	.22	.46	.74
$Median$	.3	.7	.2	.5	.8	.2	.4	.7	.3	.6	.3	.6	.2	.45	.8
$Stdev$	.1	.15	.12	.13	.11	.04	.07	.1	.1	.1	.01	.02	.02	.1	.15
N	163		160			172			186		190		187		

Table 2: Descriptive statistics of the weights.

	$Sen.$	$Mob.$	$Emo.$	$Cog.$	$Sel.$	$Pai.$	k
$mean$	.512	.449	.628	.21	.303	.407	.072
$median$	.5	.427	.65	.2	.271	.4	.036
$s.d.$	.254	.213	.198	.154	.241	.225	.108
N	60	60	60	60	60	60	60

Marital status, age and gender are the three covariates that are used to model the expected attribute levels and weights. Marital status have four levels, "single and never married", "married or living as married", "divorced or separated" and "widower". The age ranges between 19 and 84 in the sample with a median age of 52.

5.2 Analysis of the Attribute Levels

In our analysis, we used two ordered Dirichlet models for each attribute. One of the models assumed a common β, and the other had distinct precision parameters β_c for the attributes. The common precision parameter model enables us to use data from all attributes in updating parameters of single attribute utility models. We also considered the extension of the model using covariate information as discussed in Section 4. In all our analyses we used proper but diffused priors.

In Table 3 we present the summary statistics from the posterior predictive distribution of the utility at each attribute level. Note that utility across different attributes can not be compared such that we can not argue that people prefer to be at the second level of cognition rather than the second level of pain. It is neccessary to obtain the scaling constants in order to compare the utility increases in two attributes.

Table 3: Summaries for the posterior predictive distribution of the attributes' utility for the common precision ($C.P.$) and the distinct precision ($D.P.$) models.

$C.P.$	$Sensation$		$Mobility$			$Emotion$			$Cognition$		$SelfCare$		$Pain$		
	u_2	u_3	u_2	u_3	u_4	u_2	u_3	u_4	u_2	u_3	u_2	u_3	u_2	u_3	u_4
$Mean$	.341	.634	.256	.489	.747	.198	.411	.662	.339	.622	.297	.612	.22	.46	.746
$Median$	.332	.642	.242	.487	.762	.181	.406	.671	.329	.628	.284	.618	.204	.458	.761
$Stdev$	.133	.135	.122	.139	.122	.111	.137	.132	.134	.135	.129	.137	.116	.139	.122
N	20000		20000			20000			20000		20000		20000		

$D.P.$	$Sensation$		$Mobility$			$Emotion$			$Cognition$		$SelfCare$		$Pain$		
	u_2	u_3	u_2	u_3	u_4	u_2	u_3	u_4	u_2	u_3	u_2	u_3	u_2	u_3	u_4
$Mean$	.34	.636	.255	.488	.747	.198	.41	.661	.338	.624	.298	.614	.219	.458	.745
$Median$	.331	.644	.246	.489	.759	.182	.406	.67	.326	.633	.283	.623	.204	.455	.758
$Stdev$	.136	.139	.101	.115	.1	.109	.135	.131	.14	.144	.149	.159	.111	.133	.116
N	20000		20000			20000			20000		20000		20000		

In order to compare the performance of the common and distinct precision models, we used the *deviance information criterion* (*DIC*) of Spiegelhalter, Best, Carlin, and van der Linde (2002). For a generic parameter vector Θ, DIC is defined as

$$DIC = Dbar + pD, \tag{23}$$

where D $= -2log\mathcal{L}(\Theta)$, is two times the negative loglikelihood, $Dbar =$ $\text{E}_{\Theta|data}$[D] and $pD = Dbar - Dhat(\widehat{\Theta})$, where $\widehat{\Theta}$ is the posterior mean. The DIC has the general "fit $+$ complexity" form used by many model selection criteria. In (23) $Dbar$ represents the "goodness of the fit of the model where pD represents a complexity penalty as reflected by the effective number of parameters of the model. The model with the smaller

DIC is more likely to predict closer values to the observed data from a similar dataset that the likelihood is based on. In Table 4, we present the above measures for both the common and distinct precision models. Data seems to give more support to the distinct precision model.

Table 4: DIC Comparison.

	$Dhat$	$Dbar$	DIC	pD
$Common$	-2841	-2857	-2825	15.97
$Distinct$	-2909	-2930	-2888	20.99

In Figure 1 we present a comparison of the observed utility of individuals associated with the attributes cognition, self care and sensation and the posterior predictive utility.

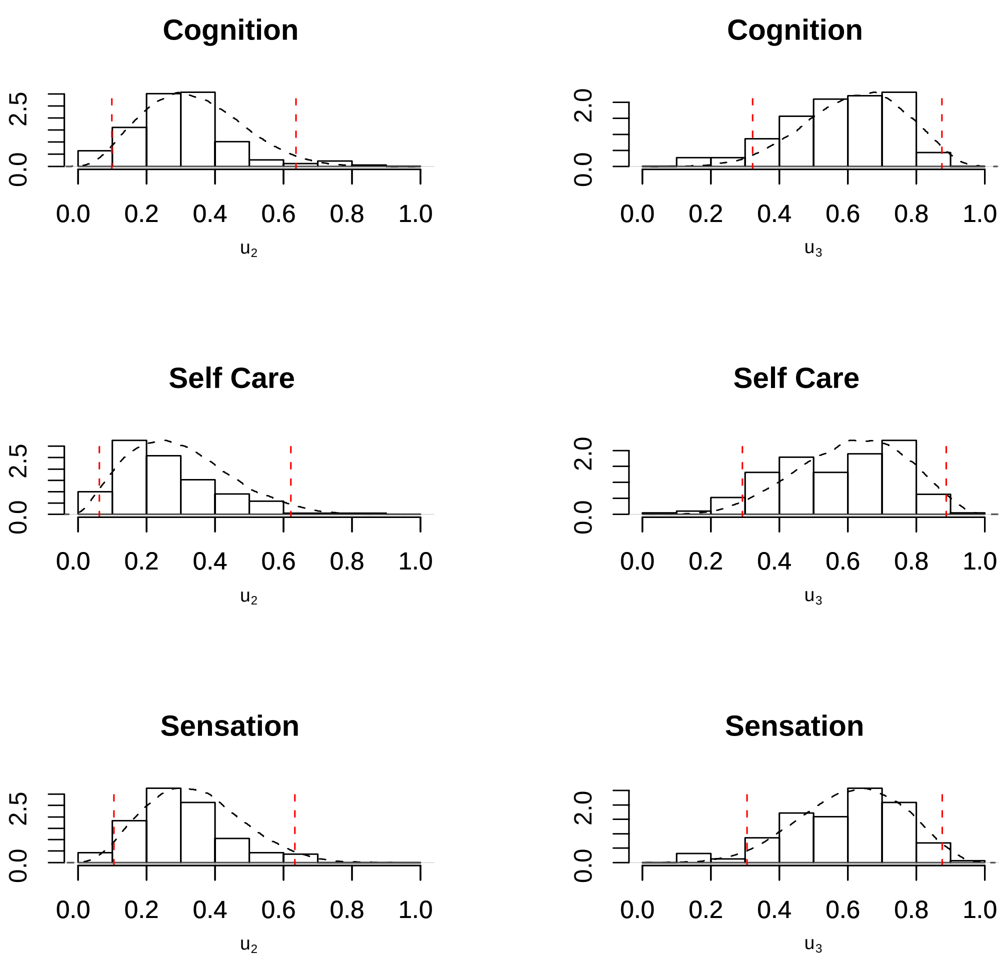

Figure 1. Histogram of the observed utility data and the plots of the posterior predictive distributon of utility of attribute levels. The dashed vertical lines are the 95% CI.

The dashed lines represent 95% credibility interval of the of the predictive utility. The figures show that the ordered Dirichlet model with distinct precision coefficients is an appropriate representation of the data. Thus, in what follows we will present our results using this model.

In Figure 2 we present the posterior predictive distributions of the utility difference between the levels of the self care attribute. It is of interest to assess the risk attitude of a random individual from the population. We can compute posterior probabilities of different risk attitudes in the population. These probabilities are presented in Table 5 for the self care attribute. We note that the probability of an individual to be risk seeking is 0.237 and risk averse is 0.106. The table compares these posterior probabilities with the relative frequencies obtained from the data.

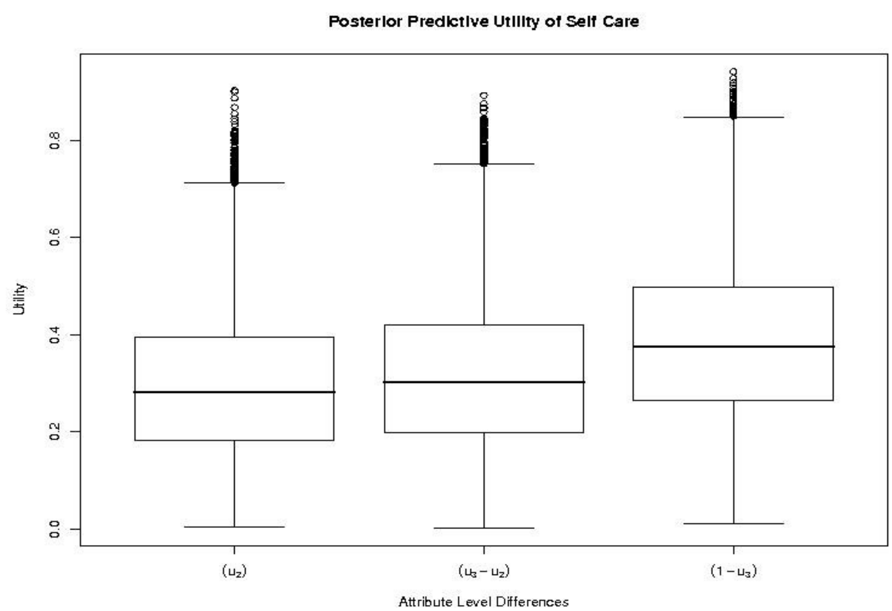

Figure 2. Posterior predictive utility differences of self care attribute.

We considered an extension of the model by including covariate information. For illustrative purposes we included the gender as the covariate in our models for each attribute. It is of interest to calculate the expected utility increase for adjacent level changes in an attribute and observe the effect of covariates. In Table 6, we present the posterior means and standard deviations of α_i's for attribute pain associated with both genders.

Table 5: Probability of different risk attitudes in self care attribute.

	Model	Data
$P(u_3 - u_2 \geq u_2, 1 - u_3 \geq u_3 - u_2)$	.237	.199
$P(u_3 - u_2 \leq u_2, 1 - u_3 \leq u_3 - u_2)$	.106	.063
$P(u_3 - u_2 \leq u_2, 1 - u_3 \geq u_3 - u_2)$	.365	.408
$P(u_3 - u_2 \geq u_2, 1 - u_3 \leq u_3 - u_2)$	.291	.33

Table 6: Expected utility increase for the levels of the attributes for different genders.

	$Sensation$		$Emotion$		$SelfCare$		$Pain$	
$Mean, s.d.$	$Male$	$Female$	$Male$	$Female$	$Male$	$Female$	$Male$	$Female$
α_2	.374, .019	.324, .013	.142, .016	.191, .01	.316, .017	.287, .013	.255, .013	.202, .009
α_3	.28, .017	.302, .012	.263, .026	.221, .01	.346, .018	.297, .013	.226, .013	.248, .01
α_4	.346, .018	.374, .013	.24, .022	.246, .011	.337, .018	.416, .014	.274, .014	.291, .011
α_5			.355, .017	.343, .012			.244, .013	.259, .01

We can probabilistically compare whether the utility of an individual with a given characteristic is greater than the utility of another individual with some other characteristic. For instance we can be interested in finding the probability of a male having a higher utility than a female at level 2 of the attribute pain. This can be accomplished by comparing the posterior distribution of the utility of the two individuals mentioned above.

$$P((u_2|Male) > (u_2|Female)) \approx$$
$$\frac{1}{S}\sum_{s=1}^{S} P\big((u_2|Male, \rho^{(s)}, \chi^{(s)}, \beta^{(s)}) > (u_2|Female, \chi^{(s)}, \beta^{(s)})\big)$$

We report the results for the pain attribute in Table 7 and observe that for worse levels, males tend to have a higher utility than females.

Table 7. Probabilistic comparison of utility

Attribute-Pain	
$P((u_2|Male) > (u_2|Female))$	.634
$P((u_3|Male) > (u_3|Female))$	.567
$P((u_4|Male) > (u_4|Female))$	.454

We also considered a random effect extension of the model with the gender covariate. In Figure 3 we illustrate the random effect of 30 individuals' ϵ_4 component of the pain attribute, sorted in relation to the mean effect. As can be seen in the diagram, there seems to be individuals, whose expected utility differs from the others when controlled for the gender effect.

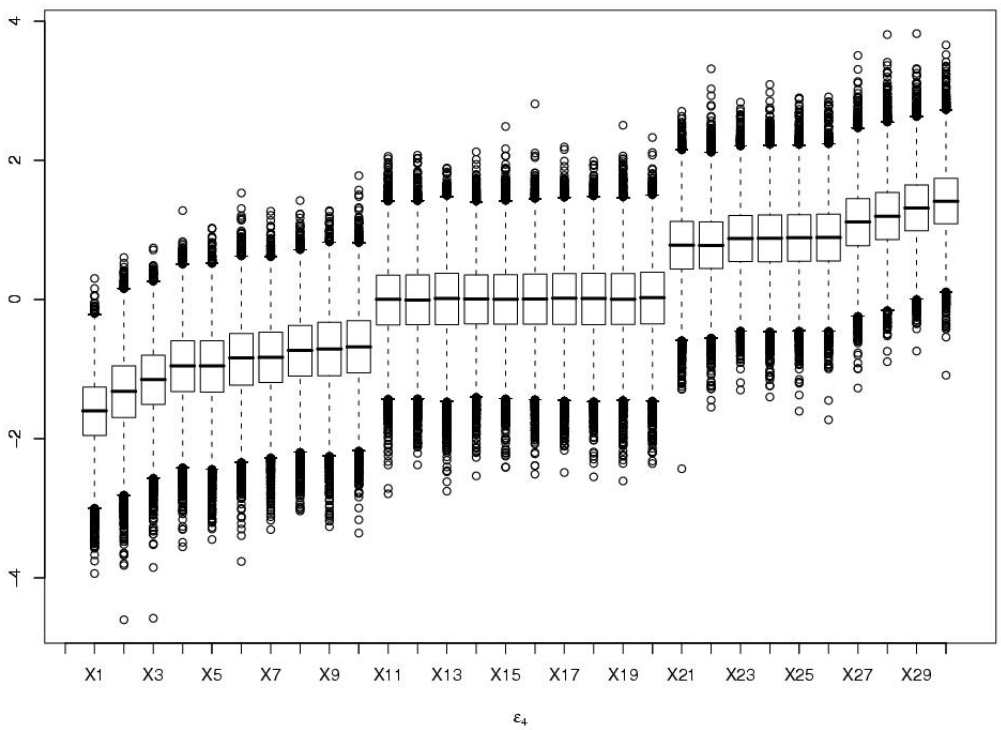

Figure 3. Random effect of 30 individuals associated with level 4 of the attribute pain controlled for gender.

5.3 Analysis of Scaling Weights

Following our development in Section 3, we present the posterior means and standard deviations of mean coefficients in table 8. The posterior mean and standard deviation of the common precision component κ is 3.957 and 0.272 respectively. The scaling constant k has a posterior predictive distribution with a mean of -0.963 and a standard deviation of .054. Thus, data strongly supports the multiplicative model rather than an additive utility model.

Table 8: Posterior mean and standard deviation of the independent beta model parameters.

$mean, sd$	$Sen.$	$Mob.$	$Emo.$	$Cog.$	$Sel.$	$Pain$
v_c	.517,.028	.457,.028	0.616,.027	0.24,.023	0.308,.025	0.41,.028

In Table 9 we illustrate the age effect on the expected weights of the attributes. As can be seen from the table, as the value of age increases expected attribute weights increase for the sensation, mobility and self care attributes. The steepest increase is observed for the sensation attribute. The other attributes, emotion, cognition and pain's weights decrease on average with age. The attribute pain has the steepest decrease of disvalue with age.

Table 9: Mean and standard deviation of the posterior predictive of scaling weights for three individuals at ages 27,55 and 81.

$mean, s.d.$	$Sen.$	$Mob.$	$Emo.$	$Cog.$	$Sel.$	$Pain$
$Age = 27$	.399,.222	.407,.222	.658,.216	.264,.2	.299,.209	.54,.226
$Age = 55$	.526,.223	.462,.223	.614,.217	.235,.19	.306,.206	.397,.216
$Age = 81$	.641,.22	.515,.229	.568,.226	.212,.18	.311,.21	.281,.203

When the model is extended by the addition of random effects the DIC significantly improves as shown in Table 10. Thus, there exists variability among individual's attribute weights which can not be captured only by covariate age.

Table 10. Model fit and complexity of the weights model.

Model	$Dbar$	$Dhat$	DIC	pD
No Covariates	-155.2	-162	-148.2	6.979
Age	-164.3	-177.3	-151.2	13.08
Age *and R.E*	-1017	-1398	-635.1	381.6

5.4 Posterior Evaluation of Health States

As discussed in Section 3 given samples from the posterior distributions of the attribute weights and utility levels, we will be able to compare the preferences between health states probabilistically. For this purpose we construct three health states $\boldsymbol{X_1} = (X_1,X_1,X_1,X_3,X_1,X_2)$, $\boldsymbol{X_2} = (X_1,X_1,X_3,X_2,X_1,X_1)$, $\boldsymbol{X_3} = (X_1,X_4,X_3,X_1,X_1,X_1)$. Note that in each case the subscript denotes the particular level of the attribute. In Figure 4 we present the boxplots associated with the posterior predictive utility distributions for the three health states.

The mean and standard deviation of $\boldsymbol{X_1}$, $\boldsymbol{X_2}$, $\boldsymbol{X_3}$ are (.619,.176), (.534,.161) and (.418,.157), respectively. Using the posterior sample values obtained from the MCMC methods we can evaluate the posterior probability that one health state is preferred to the other. This can be computed by using simulated values from posterior predictive distribution of the utility associated with each health state. There is a .658 probability that a randomly selected individual from the population prefers $\boldsymbol{X_1}$ to $\boldsymbol{X_2}$. Table 11 presents the probabilistic comparison of health states. For example, the posterior probability that $\boldsymbol{X_1}$ is preferred to $\boldsymbol{X_2}$ and $\boldsymbol{X_2}$ is preferred to $\boldsymbol{X_3}$ is 0.509 in the population.

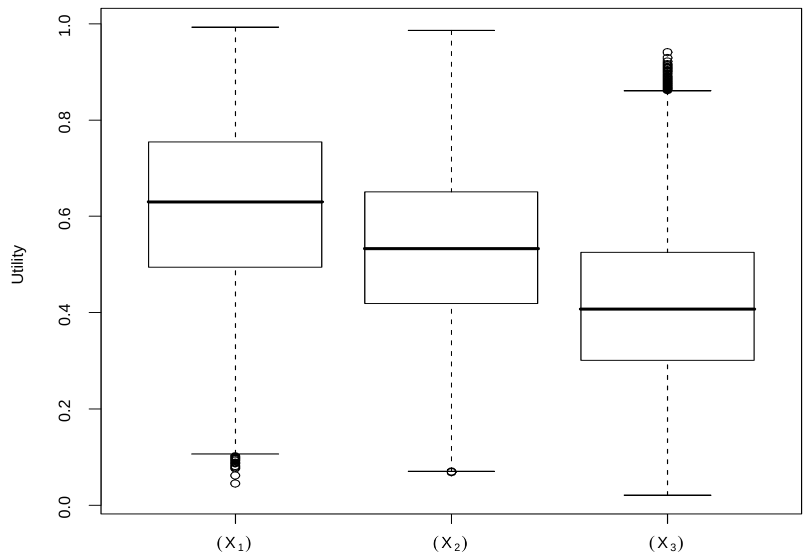

Figure 4. Posterior predictive utility distribution of $\boldsymbol{X_1}$, $\boldsymbol{X_2}$, $\boldsymbol{X_3}$.

Table 11. Posterior probabilities for health state preferences in the population.

$P(X_1 \succ X_2)$	.658	
$P(X_2 \succ X_3)$	.795	
$P(X_1 \succ X_3)$	.794	
$P(X_1 \succ X_2 \succ X_3)$	.509	
$P(X_2 \succ X_1 \succ X_3)$	.17	
$P(X_3 \succ X_1 \succ X_2)$	.033	
$P(X_1 \succ X_3 \succ X_2)$	.116	
$P(X_2 \succ X_3 \succ X_1)$	.117	
$P(X_3 \succ X_2 \succ X_1)$	.055	

When we incorporated the covariate effects on the attribute weights, these probabilities changed. In Table 12 we present the probabilistic comparison of health states using age and gender. In so doing, we only use health states $\boldsymbol{X_1}$ and $\boldsymbol{X_2}$. We note that a 27 year old male individual tends to prefer $\boldsymbol{X_2}$ to $\boldsymbol{X_1}$ whereas a female of the same age seems to be indifferent between the two states. As people of both genders get older they tend to prefer $\boldsymbol{X_1}$ to $\boldsymbol{X_2}$ however, males' tendency to prefer $\boldsymbol{X_1}$ to $\boldsymbol{X_2}$ increases more rapidly than the female.

Table 12. Probability that health state X_1 is preferred to X_2 accounting for gender and age

	$Male$			$Female$		
	$Age = 27$	$Age = 55$	$Age = 81$	$Age = 27$	$Age = 55$	$Age = 81$
$Pr(X_1 \succ X_2)$	.456	.697	.787	.515	.663	.762
$mean, s.d.(X_1)$	.482,.176	.641,.166	.723,.16	.515,.182	.622,.175	.71,.166
$mean, s.d.(X_2)$	.507,.16	.537,.16	.567,.164	.508,.159	.537,.162	.567,.164

6. Concluding Remarks

In this paper we presented a Bayesian framework for modeling and analysis of health state utility data based on MAUT models. The approach is comprehensive in that it allows for estimation of the population utilities on each dimension as well as aggregated population multiattribute utility functions and it allows for probabilistic comparison of different health states for the population.

We considered extensions of the framework via incorporation of covariate and heterogeneity effects into the analysis. We presented an implementation of the framework via Monte Carlo-based methods using actual health-state data and ilustrated what type of insights can be obtained from a Bayesian analysis of such data.

References

Brazier, J., & Deverill, M. (1999). Obtaining the Q in QALYs: A Comparison of Five Multi-attribute Utility Scales. School of Health and Related Research. The University of Sheffield.

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.

Brazier, J. (2005). Current State of the Art in Preference Based Measures of Health and Avenues for Further Research. Discussion Paper. School of Health and Related Research, The University of Sheffield.

van Dorp, J. R., Mazzuchi, T.A., & Soyer, R. (1997). Sequential inference and decision making for single mission systems development. *Journal of Statistical Planning and Inference*, 62, 207-218.

Erkanli, A., Mazzuchi, T.A., & Soyer, R. (1998). Bayesian Computations for a Class of Reliability Growth Models. *Technometrics*, 40, 14-23.

Furlong, W, Feeny, D, Torrance, G .W., & Barr, R. (2001). The Health Utilities Index (HUI®) System for Assessing Health-Related Quality of Life in Clinical Studies. McMaster University Centre for Health Economics and Policy Analysis Research Working Paper Series # 01-02.

Hazen, G. (2004). Multiattribute Structure for QALYs. *Decision Analysis*, 1, 205-216.

Kaplan, R.M., Ganiats, G.M., Sieber, W. J., & Anderson, J. P. (1988). The Quality of Well-Being Scale: critical similarities and differences with SF-36. *International Journal of Quality of Health Care*, 10, 509-520.

Keeney, R.L. (1976). A Group Preference Axiomatization with Cardinal Utility. *Management Science*, 23, 140-145.

Keeney, R. L., & Kirkwood C.W. (1975). Group Decision Making Using Cardinal Social Welfare Functions. *Management Science*, 22. 430-437.

Keeney, R.L., & Raiffa, H. (1976). Decisions with Multiple Objectives-Preferences and Value Tradeoffs, New York : Wiley.

Kharroubi, S. A., Brazier, J. E., O'Hagan, A. and Roberts, J. (2007). Modelling SF-6D health state preference data using a nonparametric Bayesian method. *Journal of Health Economics*, 26, 597-612.

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

McCabe, C., Steven, K., & Brazier, J. (2004). Utility Scores or the Health Utility Index Mark 2: An empirical assessment of alternative mapping functions. Sheffield Health Economics Group. Discussion Paper Series. Ref: 04/1.

Narayanan, A. (1991). Maximum likelihood estimation of the parameters of the Dirichlet distribution. *Applied Statistics*, 40, 365-374.

Spiegelhalter, D.J., Best, N.G., Carlin, B. P., & van der Linde A. (2002). Bayesian Measures of Model Complexity and Fit. *Journal of Royal Statistical Society B*, 64, 583-639.

Spiegelhalter, D.J., Thomas, A., Best, N., & Lunn, D. (June 2005). WinBugs User Manual (version 2.10). Medical Research Council Biostatistics Unit, Cambridge.

Torrance, G. W., Raynauld, JP, Walker, V, Goldsmith, CH, Bellamy, N. B., Band, P.A. & et al. (2002). A prospective, randomized, pragmatic, health outcomes trial evaluating the incorporation of hylan G-F 20 into the treatment paradigm for patients with knee osteoarthritis (Part 2 of 2): Economic results. *Osteoarthritis Cartilage*, 10, 518-527.

Torrance, G. W., Feeny, D.H., Furlong, W. J. , Ronald, B., Yueming, Z., & Qinan, W. (1996). Multiattribute Utility Function for a Comprehensive Health Status Classification: Health Utilities Index Mark 2. *Medical Care*, 34, 702-722.

Torrance, G. W., Furlong, W. J., Feeny, D. H., & Boyle, M. (1995). Multi-Attribute Preference Functions: Health Utilities Index. *Pharmaco Economics*, 7, 503-520.

Torrance, G. W., Boyle, M. H.& Horwood, S. P. (1982). Application of Multiattribute Utility Theory to Measure Social Utility for Health States. *Operations Research*, 30, 1043-1069.