Bayesian modeling of health state preferences: could borrowing strength from existing countries' valuations produce better estimates

Samer A Kharroubi
American University of Beirut
Beirut, Lebanon

Abstract

Background: Valuations of preference-based measure such as EQ-5D or SF6D have been conducted in different countries. There is a scope of borrowing strength from existing countries' valuations to generate better representative utility estimates. **Methods:** Data from two SF-6D valuation studies were analyzed where, using similar standard gamble protocols, values for 241 and 249 health states were devised from representative samples of the Japan and UK general adult populations, respectively. A nonparametric Bayesian model was applied to estimate a Japan value set, where the UK results were used as informative priors. Generated estimates were compared to a Japan value set estimated using Japan values alone using different prediction criterion. **Results**: The results allowed the UK data to provide significant prior information to the Japan analysis by generating better estimates than using Japan data alone. **Conclusion**: The implications of these results will be extremely crucial in countries where valuation studies are limited.

Keywords: nonparametric Bayesian methods; standard gamble; SF-6D system; preference-based health measures.

Introduction

Economic evaluation using cost-utility analysis (CUA) has increased in popularity in the health sector given the budget constraints and the increased importance of wisely allocating health resources. In order to measure the benefits from a certain treatment, CUA employs quality adjusted life years, a measure which multiplies the quality adjustment for health by the duration lived in the given health state [1]. The quality adjustment weight is generated using a preference-based measure, consisting of a classified system to describe health along with a value set denoting a utility value for every defined health state by that system. A large number of preference-based measures of health-related quality of life are currently available, including the EQ-5D [2], HUI [3], and the SF-6D [4].

There has been an increased number of datasets where preferences have been elicited for the same measure for different countries, and disparities have been accounted for. Badia et al. [5] showed quite small and potentially unimportant differences between UK, US and Spain. Johnson et al. [6] found that differences between the US and UK were potentially important. Kharroubi et al. [7-9] extended this work by developing a nonparametric Bayesian approach to model differences between these countries in a more intuitive way than the parametric random effects model of Johnson et al. [6]. Such a model brings a key potential advantage as it allows making use of findings of one country to improve those of another country, and as such generated utility estimates of the second country will be much more precise than would have been possible if that country's data was collected and analyzed on its own.

Recently, Kharroubi [10] explored this using a case study for US and UK EQ-5D data, where it is shown that the UK valuations contributed significant prior information to the US analysis. A key assumption underlying the US/UK analysis was the cultural similarities between the two countries, in addition to both having large quantity of data. However, different countries may have different preferences in addition to different population compositions, work, cultures and language. Additionally, people in different countries may have a different attitude to risk rather than health per se. All of these could potentially impact on the relative values given to different dimensions of health.

The aim of the paper is to explore the use of such a model for countries with smaller population size and with different population compositions, different types of work, greater cultural differences and different language, and for this to enable the generalizability of these approaches by making use of experience in a European country to aid the analysis of a study in another Asian country. This is explored using a case study for SF-6D Japan and UK data, where a Japan value set was estimated using the UK dataset as informative prior, and the generated estimates were compared to a Japan value set estimated using Japan values alone.

Methods

The SF-6D descriptive system: The SF-6D is a generic measure of health states. It is composed of 6 dimensions, including: *physical functioning, role limitations, social functioning, pain, mental health,* and *vitality*, each having between 4 and 6 levels. An SF-6D health state is defined by selecting one statement from each dimension, starting with physical functioning and ending with vitality. Level 1 in each dimension represents no loss of health or functioning in that dimension, so that state 111111 denotes perfect health. The worst possible state is 645655, known as 'the pits'. A total of 18000 health states can be defined in this way.

The valuation survey and data set

UK: A sample of 249 health states defined by the SF-6D was valued by a representative sample of the UK general population (n = 836). A detailed description of the selection methods of respondents and health states is reported elsewhere [4]. Each respondent was asked to rank, and then value, six of these states using the standard gamble (SG) technique. The SG technique asked the respondents to value five of the six SF-6D states against the perfect health and the "pits". Respondents were then asked in the sixth SG question to value 'pits'. Depending on whether they thought this state was better or worse than death they would be asked to consider one of the following prospects: (i) the certain prospect of being in the "pits" state and the uncertain prospect of full health or immediate death; or (ii) the certain prospect of death and the uncertain prospect of full health or the "pits" state [5]. The chances of the best outcome occurring is varied until the respondent is indifferent between the certain and uncertain prospects. Any negative value was bounded to -1 and it referred to worse than death. The other states were chained onto the scale 1 to 0, with 1 being the perfect health state and 0 indicating death. These adjusted SG values form the dependent variable (y) in the model discussed below.

Of the original 836 respondents, a total of 225 respondents were excluded for many reasons. For example, 130 respondents failed to value the "pits" state, making it impossible to generate an adjusted utility value. From the remaining 611 included respondents, there were 148 missing values from 117 respondents resulting in 3518 observed SG valuations across the 249 health states. The details of the valuations for each of the 249 SF-6D UK health states can be found in [4].

Japan: The Japan study comprised a sample of 241 health states, which were selected and valued according to the UK selection and valuation procedures [9]. Due to the necessity of introducing some modifications to the Japanese version of the SF-6D, it was not possible to include the same health states as in the UK study. Each respondent was asked to rank and value seven health states, and the interview protocol was modelled on that used in the UK study. Out of the original 600 respondents, a total of 135 respondents were excluded from the analysis. Furthermore, similar exclusion criteria were applied to ensure comparability with the UK study [4]. From the remaining 465 individuals included in the study there were 185 missing, resulting in 3070 observed SG valuations across 241 health states. A detailed valuation of the 241 SF-6D Japan health states is available in [11].

Modelling

Kharroubi [10] proposed the following model

$$y_{ij} = 1 - \alpha_j\{1 - u(\mathbf{x}_{ij})\} + \varepsilon_{ij} \qquad \textbf{(1)}$$

where, for i=1,2,…,I_j and j=1,2,…,J, x_{ij} is the i^{th} health state valued by the respondent j in the Japan study, y_{ij} is the respondent j's SG valuation for the given health state i, α_j is random individual effect and ε_{ij} is a random error term. Let $\boldsymbol{t}_j$, be the individual characteristics of respondent j, the following distribution has been suggested by Kharroubi [10]:

$$\alpha_j \sim LN\left(t_j^T\gamma, \tau^2\right) \text{ and } \quad \varepsilon_{ij} \sim N(0, v^2).$$

where $\boldsymbol{\gamma}$ is the vector of coefficients for the covariates and τ^2 and v^2 are the parameters to be further estimated. Furthermore, we define $\boldsymbol{u}(\mathbf{x})$ and $u_{UK}(\mathbf{x})$ as the utility functions for health state $\mathbf{x}$ valued in the Japan and the UK studies respectively, for which Kharroubi [10] modeled the prior distribution for $\boldsymbol{u}(\mathbf{x})$ to have a multivariate normal distribution with mean defined as

$$E\big(u(\mathbf{x})\big) = E(u_{UK}(\mathbf{x})) + \gamma + \beta'\mathbf{x} \qquad \textbf{(2)}$$

and variance-covariance matrix

$$\text{cov}(u_{UK}(\mathbf{x}), u_{UK}(\mathbf{x}')) + \sigma^2 \text{c}(\mathbf{x}, \mathbf{x}') \qquad \textbf{(3)}$$

where $E\big(u_{UK}(\mathbf{x})\big)$ and $\text{cov}(u_{UK}(\mathbf{x}), u_{UK}(\mathbf{x}'))$ represent the expected utility value for health state x and the variance-covariance matrix between $u_{UK}(\mathbf{x})$ and $u_{UK}(\mathbf{x}')$ for two different states $\mathbf{x}$ and $\mathbf{x}'$ in the UK study and they are both obtained from the analysis of the existing UK data. Notice that the inclusion of $E\big(u_{UK}(\mathbf{x})\big)$ and $\text{cov}(u_{UK}(\mathbf{x}), u_{UK}(\mathbf{x}'))$ in equations (2) and (3) allows for the existing evidence from the UK study to contribute as informative prior to the Japan population utility function $\boldsymbol{u}(\mathbf{x})$. For further details, refer to Kharroubi [10].

Given Equations (2) and (3), notice that $\mathbf{x}$ is a vector consisting of discrete levels, each from one of the six dimensions, with γ, β and σ^2 being unknown parameters to be estimated. Notice also that the mean function of $\boldsymbol{u}(\mathbf{x})$ represents a belief that the utility will be roughly additive linear function of $\mathbf{x}'$s levels in each dimension. Furthermore, the actual function is free to vary around this mean according to its multivariate normal distribution, and so it may take absolutely any form. Furthermore, when $\mathbf{x}$ and $\mathbf{x}'$ are close enough, $u(\mathbf{x})$ and $u(\mathbf{x}')$ seem to have high correlation $\text{c}(\mathbf{x}, \mathbf{x}')$, expressed as

$$\text{c}(\mathbf{x}, \mathbf{x}') = \exp\{-\textstyle\sum b_d (x_d - x'_d)^2\} \qquad \textbf{(4)}$$

where for d = 1, 2, …, 6, x_d and x'_d are the levels of dimension d in the health states $\mathbf{x}$ and $\mathbf{x}'$, respectively, and b_d is a roughness parameter that controls the degree of adherence of the true utility function to a linear form in dimension d. Kharroubi [12] provides a more through explanation about this specific point.

Finally, note that the population mean health state utility in (1) 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 $u(\mathbf{x})$. See Kharroubi [12] for more details on the computation of $\bar{\alpha}$.

General theory and full technical details of the Bayesian model presented here are discussed in Kharroubi [10]. Programs to undertake the Bayesian model were written in Matlab and codes are available upon request. The Matlab codes are not general and the user will need to modify them for his/her own purposes.

Results

For this analysis, the nonparametric Bayesian model was applied to estimate a Japan value set, where the UK results were used as informative priors (to be indicated by Japan/UK model hereinafter), and the estimates were compared to the estimates generated from analyzing the Japan data excluding the UK data (to be indicated by Japan model).

The models are compared in terms of their predictive ability in Figure 1, which shows the Japan predicted (line marked with squares) and actual (line marked with diamonds) mean valuations for the 241 health states valued in the survey with health states ordered by actual health state values, along with their computed errors (line marked with triangles). Figure 1a displays the results obtained from the Japan model; whereas Figure 1b shows the corresponding results from Japan/UK model. It is apparent from these plots that the utilities obtained from Japan/UK model for various SF-6D health states have higher accuracy than those estimated from the Japan model. Furthermore, these plots reveal that the Japan model tends to under predict at high health states and over predict at poor health states, which is not the case with Japan/UK model. Moreover, the predictions tend to have a smaller variation in Japan/UK model, which is reflected in the smaller peaks of the computed error, as well as from the closer position of the diamonds-line and squares-line (Figure 1b), thus, reducing the possibility of systematic bias for the Japan/UK model compared to the Japan model. Finally, across the 241 states that were used in the study, the Japan/UK model has a better predictive performance with a RMSE of 0.076 versus 0.096 for the Japan model.

A final comparison of the two models is to conduct an out-of-sample leave-one out prediction at the level of health states. Data relating to 12 selected health states were removed randomly and sequentially from both sets of data, and the two models were fitted on the reduced data of 240 health states. The Q-Q plots of standardized predictive errors

for the 12 out of sample health states are presented in Figure 2a for the Japan model and Figure 2b for Japan/UK model. The straight line in both figures represents a reference line passing through the first and fourth quartiles, which is helpful for judging whether the points are linear. As observed, points in Figure 2b are quasi aligned on the reference line compared to those in Figure 2a, validating our hypothesis that Japan/UK model has better predictive abilities than the Japan model.

Discussion

Here we have applied a nonparametric Bayesian model to the existing Japan–UK SF-6D valuations in the aim of checking whether borrowing extra strength from the UK data allows the generation of better estimates of Japan utilities than analyzing its data separately. The findings proved that using Japan data alongside the already existing UK data produced Japan utility estimates better than using the Japan study data alone for all prediction criterion used, including predicted versus actual mean health state valuations, mean predicted error, RMSE and an out-of-sample prediction.

Experimental studies for deriving health state values like the EQ-5D, HUI or SF-6D need to be conducted in different countries and the evidence emerging is that there are differences in their values. However, such work is costly and is potentially wasteful. The work presented here suggests how making use of the already existing data as substantial prior information improve the accuracy of prediction. This offers the potential of reducing the number of states to be valued and, as such, reducing the cost of cross-country valuation. Such an analysis will be hugely important in countries lacking the ability to conduct large evaluation studies. To our knowledge, this potential benefit has not been explored previously, and forms an important research agenda for the future.

Our basic models equation (1) offers a major added advantage: it has the potential to allow for more than two countries to be analysed. Additionally, equations (2) and (3) may be generalized to handle multi-countries. Indeed, equations (2) and (3) can be generalized further to generic forms

$$E\big(u(\mathbf{x})\big) = \sum_{k=1}^{n} E(u_k(\mathbf{x})) \; + \gamma + \beta'\mathbf{x}$$

and variance-covariance matrix

$$\sum_{\boldsymbol{k=1}}^{\boldsymbol{n}} \mathrm{cov}(u_k(\mathbf{x}), u_k(\mathbf{x}')) + \sigma^2 \mathrm{c}(\mathbf{x}, \mathbf{x}')$$

where $\sum_{k=1}^{n} E(u_k(\mathbf{x}))$ is the total mean utility of health state x and $\sum_{\boldsymbol{k=1}}^{\boldsymbol{n}} \mathrm{cov}(u_k(\mathbf{x}), u_k(\mathbf{x}'))$ is the total variance-covariance matrix between $u_k(\mathbf{x})$ and $u_k(\mathbf{x}')$ for two different states $\mathbf{x}$ and $\mathbf{x}'$, all of which are readily available from the analysis of the *n* available countries data. Work is in progress on demonstrating this idea for three countries.

References

1- Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). Methods for the economic evaluation of health care programmes.
2- Brooks, R., & Group, E. (1996). EuroQol: the current state of play. *Health policy*, *37*(1), 53-72.
3- Feeny, D., Furlong, W., Torrance, G. W., Goldsmith, C. H., Zhu, Z., DePauw, S., ... & Boyle, M. (2002). Multiattribute and single-attribute utility functions for the health utilities index mark 3 system. *Medical care*, *40*(2), 113-128.
4- Brazier, J., Roberts, J., & Deverill, M. (2002). The estimation of a preference-based measure of health from the SF-36. *Journal of health economics*, *21*(2), 271-292.
5- Badia, X., Roset, M., Herdman, M., & Kind, P. (2001). A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. *Medical Decision Making 21*(1), 7-16
6- Johnson, J. A., Luo, N., Shaw, J. W., Kind, P., & Coons, S. J. (2005). Valuations of EQ-5D health states: are the United States and United Kingdom different?. *Medical care*, *43*, 221-228.
7- Kharroubi, S. A., O'Hagan, A., & Brazier, J. E. (2010). A comparison of United States and United Kingdom EQ-5D health states valuations using a nonparametric Bayesian method. *Statistics in medicine*, *29*(15), 1622-1634.
8- Kharroubi SA, Brazier J, McGhee S. A comparison of Hong Kong and United Kingdom SF-6D health states valuations using a non-parametric Bayesian method. Value in Health. 2014; 17(4): 397-405.
9- Kharroubi SA. A comparison of Japan and United Kingdom SF-6D health states valuations using a non-parametric Bayesian method. Applied Health Economics Policy. 2015; 13: 409-20.
10- Kharroubi, SA (2018). Valuations of EQ-5D health states: could United Kingdom results be used as informative priors for United States. *Journal of Applied Statistics;* 45(9), pp. 1579-1594
11- Brazier, J. E., Fukuhara, S., Roberts, J., Kharroubi, S., Yamamoto, Y., Ikeda, S., ... & Kurokawa, K. (2009). Estimating a preference-based index from the Japanese SF-36. *Journal of Clinical Epidemiology*, *62*(12), 1323-1331.
12- Kharroubi, S. A., O'Hagan, A., & Brazier, J. E. (2005). Estimating utilities from individual health preference data: a nonparametric Bayesian method. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *54*(5), 879-895.

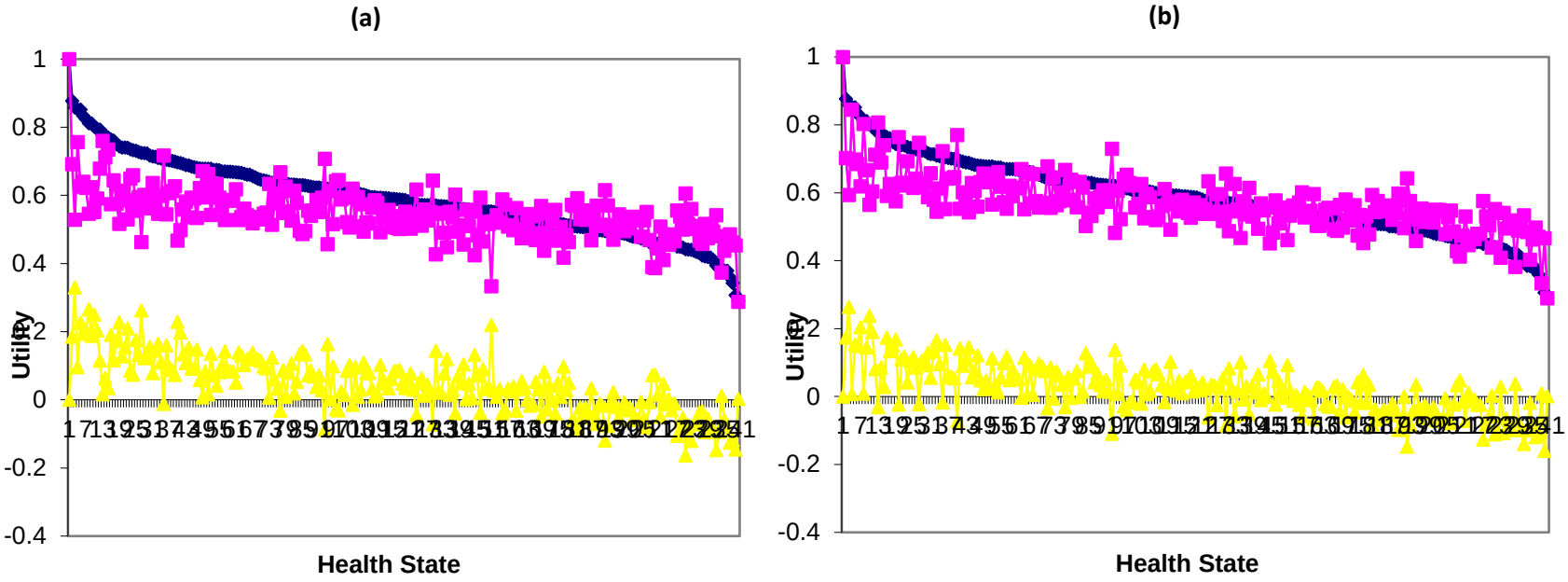

Fig. 1. Sample mean and predicted health state valuations for (a) the Japan model and (b) Japan/UK model

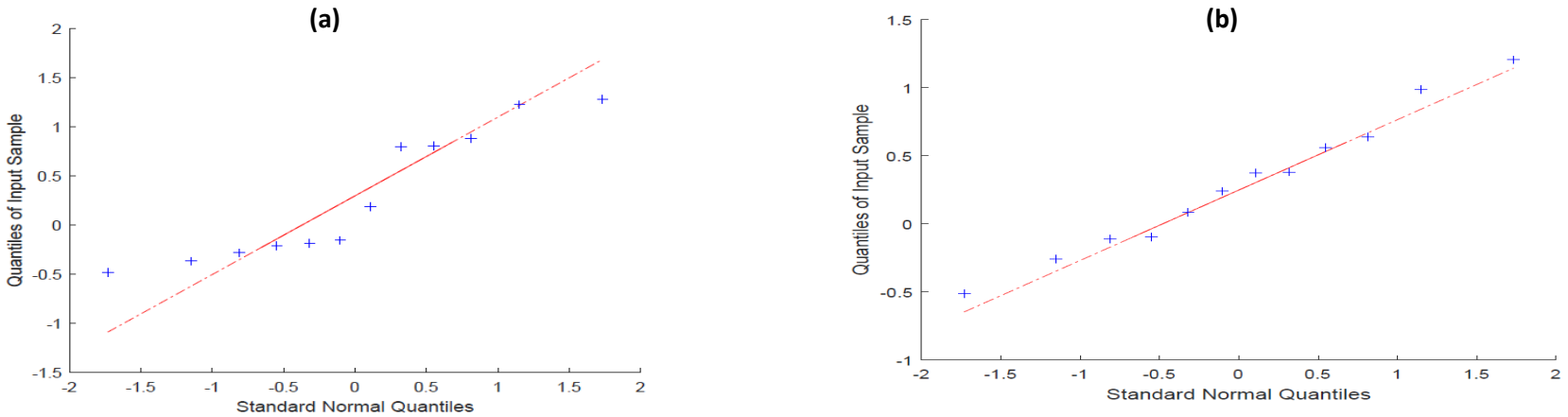

Fig. 2. Q-Q Plot for the 12 out of sample health states for (a) the Japan model and (b) Japan/UK model.