Estimating health state utility values (QALY weights) from discrete choice experiments – a QALY space model approach

Yuanyuan Gu, Richard Norman, Rosalie Viney
Centre for Health Economics Research and Evaluation (CHERE), University of Technology Sydney (UTS)

International Choice Modelling Conference 2013
Outline

- Background
  - “health state utility values (QALY weights)”

- Motivation/Objective
  - Previous methods: an average person’s QALY weights
  - Our methods: population mean QALY weights (used in economic valuation literature)

- Methods
  - Based on Mixed Logit (MIXL)
  - “preference space” vs. “QALY space”

- Results and discussion
Background: some basics

- **QALY and QALY weights**
  - Quality Adjusted Life Years - outcome measure considering both the quality and quantity of life lived
  - The basis in Cost-Utility Analysis (CUA).
  - CUA: \( \text{ICER} = \frac{\text{Cost increment}}{\text{QALY increment}} \) (to compare multiple health programs)
  - Number of QALYs =
    - Utility score associated with a health state (QALY weights) \( \times \) Number of years lived in this health state
  - Utility value for perfect health = 1;
  - Utility value for death = 0
How to estimate utility value of a health state?

- “Health state” – measuring health
- “Utility value” – valuing health

Health measure instruments

- EQ-5D, SF-6D, AQOL (assessment of QoL), ...

Valuation techniques

- Time Trade Off (TTO), Standard Gamble (SG), Discrete Choice Experiments (DCEs), ...

Yuanyuan Gu
Background: measure health using the EQ-5D

<table>
<thead>
<tr>
<th>Mobility</th>
<th>Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>I have no problems in walking about</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>I have some problems in walking about</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>I am confined to bed</td>
</tr>
<tr>
<td>Self Care</td>
<td>1</td>
<td>I have no problems with self-care</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>I have some problems washing and dressing myself</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>I am unable to wash and dress myself</td>
</tr>
<tr>
<td>Usual Activities</td>
<td>1</td>
<td>I have no problems with performing my usual activities</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>I have some problems with performing my usual activities</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>I am unable to perform my usual activities</td>
</tr>
<tr>
<td>Pain / Discomfort</td>
<td>1</td>
<td>I have no pain or discomfort</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>I have moderate pain or discomfort</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>I have extreme pain or discomfort</td>
</tr>
<tr>
<td>Anxiety / Depression</td>
<td>1</td>
<td>I am not anxious or depressed</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>I am moderately anxious or depressed</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>I am extremely anxious or depressed</td>
</tr>
</tbody>
</table>

- Five domains – mobility (MO), self-care (SC), usual activities (UA), pain/discomfort (PD), and anxiety/depression (AD)
- Three levels in each domain – level 1 is the best and level 3 is the worst
- In total 243 health states
Background: valuation using TTO

- Describing health state using a vector of dummy variables
  - $X = (MO2, MO3, SC2, SC3, UA2, UA3, PD2, PD3, AD2, AD3)$
  - Perfect health: (0,0,0,0,0,0,0,0,0,0); Worst health: (0,1,0,1,0,1,0,1,0,1)

- The algorithm
  - Utility value $= 1 + X\beta$ where $\beta$ represents “utility decrements”
  - The UK algorithm also considers an interaction term in $X$

- Valuation using TTO
  - Select a sample of EQ-5D health states and use TTO to elicit respondents’ utility scores for these health states
  - Estimate utility decrements $\beta$ by regressing these scores on the sampled health states
  - Predict the utility values (QALY weights) for all EQ-5D health states by using the estimated decrements
The DCE approach: literature

- Limitations of TTO/SG
  - Cognitively challenging
  - Tendency for scores to fall at extreme values (Norman et al, 2010)

- The DCE literature
  - Design and estimation: e.g., Flynn et al, 2010 (PE); Bansback et al, 2012 (JHE); Viney et al, 2013 (HE)
  - Distributional weights using characteristics: Lancsar et al, 2011 (JHE)
  - Non-linear time effect: Viney et al, 2013 (HE)
  - Compare DCE with TTO and other approaches: Bansback et al, 2012 (JHE); Viney et al, 2013 (HE); Brazier et al, 2013 (EJHE)
The DCE approach: design

The design for the EQ-5D DCE

- Orthogonal design
- 6 attributes: an EQ-5D health state (MO, SC, UA, PD and AD; three levels: no problem, moderate problem, and serious problem) and the number of years to live in that state (5 levels: 1, 2, 4, 8 and 16 years)
- 3 alternatives (unlabelled): a pair of scenarios and an “immediate death” option (complete ranking but only use the ranking between two scenarios)
- 15 choice sets
- A total of 1,031 individuals completed the survey
The DCE approach: an example

If you had to choose between the following scenarios:

**Scenario A**
- You have no problem in walking about
- You have some problems washing and dressing yourself
- You have no problems with performing your usual activities
- You have moderate pain or discomfort
- You are extremely anxious or depressed

You will live in this state for 4 years, then die.

**Scenario B**
- You have some problems in walking about
- You have no problems with self-care
- You are unable to perform your usual activities
- You have extreme pain or discomfort
- You are not anxious or depressed

You will live in this state for 4 years, then die.

**Scenario C**
- Death

Which is best? Which is worst?
The DCE approach: utility function

- **Defining the latent utility function**
  - Assumed utility functions for individual $i$ of option $j$ in choice set $s$:

$$U_{isj} = \alpha TIME_{isj} + \beta X'_{isj} TIME_{isj} + \epsilon_{isj}$$

where $X'_{isj}$ represents a set of dummy variables relating to the levels of the EQ-5D health states

- **Why this?**
  - The zero condition – for a duration of zero life years, all quality of life levels are equivalent (Bleichrodt et al., 1997).
  - The time effect is linear – TTO/SG assumption
The DCE approach: normalisation (from latent utility space to health utility scale)

- Several normalisation methods (Bansback et al, 2012; Viney et al, 2013)
- The main idea is that the utility value of a health state is its marginal utility of $TIME$ on the latent scale, i.e.,

$$\frac{\delta U}{\delta TIME} = \alpha + \beta X'$$

In the case of full health, its marginal utility of $TIME$ on the latent scale is

$$\frac{\delta U}{\delta TIME} = \alpha$$

which needs to be normalised to be 1 under the QALY model. Hence the normalising constant is $\alpha$. 

Yuanyuan Gu
The DCE approach: normalisation (from latent utility space to health utility scale)

- Under the DCE frame the algorithm is

  \[
  \text{Utility value of } X = 1 + \frac{\beta}{\alpha} X'
  \]

- Utility decrements = \( \frac{\beta}{\alpha} \)

- Following the TTO approach we need to estimate the population mean utility decrements.
Motivation

Current estimation methods:

1. Estimate $\hat{\alpha}$ and $\hat{\beta}$ using conditional logit (Bansback et al, 2012) or random effects probit (Viney et al, 2013)

2. Estimate utility decrements by computing $\frac{\hat{\beta}}{\hat{\alpha}}$

What is $\frac{\hat{\beta}}{\hat{\alpha}}$?

- It is not the population mean estimate of $\frac{\beta}{\alpha}$ although $\hat{\alpha}$ and $\hat{\beta}$ are population mean estimates of $\alpha$ and $\beta$.

- Imagine there is a person whose $\alpha$ and $\beta$ are the same as $\hat{\alpha}$ and $\hat{\beta}$. We call this person “the average person” and $\frac{\hat{\beta}}{\hat{\alpha}}$ may be seen as the average person’s utility decrements.
Motivation

- “Average valuation over whole population” vs. “An average person’s valuation”
  - Conceptually different approaches and different interpretations in policy valuation
  - TTO/SG framework: average valuation
  - DCE framework: an average person’s valuation (current methods)

- “Average valuation” under the DCE framework?
Mixed logit (MIXL): preference space

- In the MIXL the coefficients are assumed to be random: $\alpha_i$ and $\beta_i$, implying that each person has a different taste.

- The framework for estimating population mean utility decrements:
  - Specify the distributions of $\alpha_i$ and $\beta_i$
  - Estimate the distribution parameters using the sample
  - Derive the population distributions of $\beta_i/\alpha_i$
  - Find the means of the population distributions
Mixed logit: $\alpha_i, \beta_i \sim ?$ distribution

- **Normal densities? No, because**
  - $\beta_i / \alpha_i$ should have finite moments
  - $\alpha_i$ represents a person’s preference for the life duration at perfect health. The longer life the better, so $\alpha_i > 0$.
  - $\beta_i$ represents a person’s utility loss when his/her health condition deteriorates from the perfect level to a worse level, so $\beta_i < 0$.

- **Bounded distributions (Log-normal)**
  - Need to change the signs of the data corresponding to $\beta_i$ to their opposite
Challenges of using preference space

- Distribution of the ratio of two random variables
  - A longstanding problem
  - Particularly investigated in the WTP literature

- Major challenges
  - Extreme values arising from the reciprocal of a random variable – mean estimates are affected
  - The distribution of $\beta_i / \alpha_i$ is induced from our assumptions on the distributions of $\beta_i$ and $\alpha_i$. 
The QALY space model

- Adapted from WTP literature
  - WTP space model (Train and Week, 2004)
  - Reparameterisation of the preference space model

- Preference space model:
  \[ U_{isj} = \alpha_i TIME_{isj} + \beta_i X_{isj} TIME_{isj} + \varepsilon_{isj} \]

QALY space model:
\[ U_{isj} = \alpha_i (TIME_{isj} + \gamma_i X_{isj} TIME_{isj}) + \varepsilon_{isj} \]

\[ \gamma_i = \beta_i / \alpha_i \] is the utility decrement
Three models

- **M1 (Preference space):**
  \[ \alpha_i \text{ and } \beta_i \text{ are both log-normal} \]

- **M2 (QALY space):**
  \[ \alpha_i \text{ and } \gamma_i \text{ are both log-normal} \]

- **M3 (QALY space):**
  \[ \alpha_i \text{ is log-normal and } \gamma_i \text{ is Johnson’s SB} \]
  i.e., \[ \gamma = K \exp(\varepsilon)/(1 + \exp(\varepsilon)) \] where \( \varepsilon \) is normal
  \( \gamma \) is bounded between 0 and \( K \)
Johnson’s SB distribution

- First introduced into the choice modelling literature by Kenneth Train to replace log-normal distribution because
  - Also a bounded distribution
  - Much thinner tails

Other nice natures:

- A wide variety of distributions such as normal, log-normal, Weibull, and modified beta can be satisfactorily fitted by the Johnson’s SB distribution.
- It can accommodate data with two modes spiked at the lower and upper bounds.
Estimation

- Bayesian
- All random parameters are fully correlated
- M3 (Johnson’s SB):
  - Difficult to estimate because the bounds and the variance are correlated – model underidentified.
  - Typical approach: grid search (fixing the bound parameter to a series of numbers and compare) – ok for the univariate case but here we need to estimate a multivariate distribution with 10 dimensions.
  - Our approach: estimate the bound parameters together with other parameters simultaneously by using informative prior distributions on the bounds – to construct Bayesian identifiability.
Comparing the mean utility decrement estimates from different models

<table>
<thead>
<tr>
<th></th>
<th>clogit</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
</tr>
</thead>
<tbody>
<tr>
<td>MO2</td>
<td>-0.12</td>
<td>-0.13</td>
<td>-0.13</td>
<td>-0.14</td>
</tr>
<tr>
<td>MO3</td>
<td>-0.52</td>
<td>-0.77</td>
<td>-0.76</td>
<td>-0.63</td>
</tr>
<tr>
<td>SC2</td>
<td>-0.12</td>
<td>-0.15</td>
<td>-0.15</td>
<td>-0.16</td>
</tr>
<tr>
<td>SC3</td>
<td>-0.29</td>
<td>-0.42</td>
<td>-0.42</td>
<td>-0.36</td>
</tr>
<tr>
<td>UA2</td>
<td>-0.10</td>
<td>-0.10</td>
<td>-0.10</td>
<td>-0.09</td>
</tr>
<tr>
<td>UA3</td>
<td>-0.19</td>
<td>-0.25</td>
<td>-0.24</td>
<td>-0.23</td>
</tr>
<tr>
<td>PD2</td>
<td>-0.11</td>
<td>-0.11</td>
<td>-0.11</td>
<td>-0.11</td>
</tr>
<tr>
<td>PD3</td>
<td>-0.50</td>
<td>-0.73</td>
<td>-0.72</td>
<td>-0.59</td>
</tr>
<tr>
<td>AD2</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.14</td>
<td>-0.14</td>
</tr>
<tr>
<td>AD3</td>
<td>-0.37</td>
<td>-0.55</td>
<td>-0.54</td>
<td>-0.47</td>
</tr>
<tr>
<td>AIC</td>
<td>17862</td>
<td>15250</td>
<td>15244</td>
<td>15170</td>
</tr>
</tbody>
</table>

M1: MIXL using preference space (decrement = log-normal/log-normal)
M2: MIXL using QALY space (decrement = log-normal)
M3: MIXL using QALY space (decrement = Johnson’s SB)
Utility decrements’ distributions estimated from two QALY space models

Yuanyuan Gu
Comparing the estimated health state utility values from different models (all 243 states)

Best

Worst
Thank you!