

## **Author's response to reviews**

**Title:** The Episodic Random Utility Model Unifies Time Trade-Off and Discrete Choice Approaches in Health State Valuation

### **Authors:**

Benjamin M Craig ([benjamin.craig@moffitt.org](mailto:benjamin.craig@moffitt.org))

Jan J V Busschbach ([j.vanbusschbach@erasmusmc.nl](mailto:j.vanbusschbach@erasmusmc.nl))

**Version:** 2 **Date:** 12 November 2008

### **Author's response to reviews:**

Thank you very much for your thorough reviews of my paper, MS: 6995120382048266, "The Episodic Random Utility Model Unifies Worse than Death and Better than Death TTO Responses in Health State Valuation." We made substantial changes in response to your comments, and feel that the manuscript is much improved. In particular, we

Below please find detailed responses to each of the comments made.

Editor comments:

In particular, it would improve the accessibility of the paper if additional introductory material could be provided to summarize:

- The conceptual framework of TTO valuation
- The conceptual framework of DCE
- The theoretical umbrella perspective to be described by lay terms
- The intuitive plausibility of the episode approach

The manuscript has been revised to improve the accessibility of the paper by incorporating feedback from the EuroQoL Group members and the faculty at University of Sheffield. The following paragraphs are particularly noteworthy:

1. Page 4, paragraph 2 introduces the theoretical umbrella in lay terms with an illustrative figure.
2. Page 5, paragraph 3, compares the issues with worse than death responses to outliers in cost data. Each threaten the face validity of estimates.
3. Page 10, paragraph 2, compares the issue of ratio statistics to the estimation of ICERs, which is commonly understood and persuasive.
4. Page 10, paragraph 3, compares the coefficient approach to response error, a standard component of measurement theory that appeals to psychometricians.
5. Page 10, previously introduce formula concerning minor aspects of equation 4, which we felt distracted the reader from the central objectives of the paper. These discussions were removed from the manuscript.

Overall, the manuscript was edited to broaden the readership based on feedback.

Reviewer 1:

1. The authors used only a single rather well-behaved case study on which to validate their method. In other words, their results are specific to the seminal United Kingdom Measurement and Valuation in Health (MVH) study. It would be useful to get a feeling for the generalisability of the validation results to other scenarios.

Yes, estimates based on other countries are currently in progress, specifically the US and Spain. The UK MVH study is the most well-known, which makes it the best platform to introduce this model and its two estimators.

2. The results seem eminently sensible but the authors do not provide any code (software) to carry out the analysis. Could this be done in, say, Excel, and if not is there a simpler but adequate version that could be?

Yes, the coefficient approach and the mean slope approach could be done in Excel. As shown in equation 4, each is a summation of responses with some slight manipulation. The exploded probit requires knowledge of maximum likelihood estimation and requires a deeper understanding of both econometrics and statistical software.

3. The authors develop and validate an episodic random utility model (RUM) that remedies concerns with worse than death (WTD) responses. This is done well. It would be helpful to get a feeling for the performance of the model of interest as far as model checking is concerned. Apart from what's already been presented here, how do make sure that the model of interest is good enough, and if so, how?

In this paper, we estimate values for hypothesized states using a coefficient approach. This approach does not require any maximization (MLE or OLS), because a coefficient is simply a mean over a mean. In future work, we will use a multi-attribute utility function, similar to Dolan (1997). Under such a regression framework, R2 may be an appropriate measure, but this would ignore individual clustering.

4. The authors illustrate the relationship between the episodic and instant RUM predictions via Figure 1. It would be of interest to consider a table for the prediction means and show the strength of the proposed model via different statistical techniques, Mean squared error for example. Also, it would be worth checking the model and comparing it with the instant one as far as monotonicity (where some better states are assigned a lower value than worse states) is concerned.

Agreed. Many of those who attend my talks requested these estimates, which are now included as an appendix to this paper.

The episodic model attempts to fit the value of the episode and the instant model attempts to fit the value of the instant; therefore, the differences in purposes prevent the use of mean squared error.

5. It would be interesting to see an extension of their model to cover more possibilities. The authors mentioned the case of reestimating each country-specific valuation sets using the episodic RUM and to further examine duration effects in components and errors...Any more cases?

More work is currently being conducted. Specifically, value sets for the UK, US, the Netherlands, and Spain.

6. Discussion needs more attention as far as highlighting findings, results and future work is concerned.

The discussion has been heavily revised to given more attention to the central findings of this paper.

Minor Comments:

1. p2, heteroscedasticity instead of heteroskedasticity

In the United States, we typically use heteroskedasticity. For example, Arthur Goldberger's book, "A Course in Econometrics" was the central text for many years, and it uses this spelling. I am deeply indebted to Dr. Goldberger, who introduced me to the field of econometrics.

2. p5, right after equation 5. whats  $\bar{d}_1$

Equation 5 was dropped from the manuscript.

3. p6 equation 6 is really hard to follow!

Equation 6 describes how ties are incorporated in the likelihood function under the Efron approach.

4. p8, 3rd paragraph. How do you get the results?

The addition of the appendix allows the reader to see the confidence intervals as well as other variations in estimates.

Reviewer 2:

The article will be quoted a lot but I think it is a bit defensive. The episode approach random utility model is presented as an arbitrary choice and is then in the end preferred because the results are better than the results of the instant model. However I would strongly recommend that the authors consider adding text which defends that the episode model is better on principle!!

Agreed. As mentioned in the response to the editor, we added comparisons to ICERs, psychometrics, and cost analyses that demonstrate how the coefficient approach is preferred on principle.

