HESG # P36

Weber's QALY: TTO and DCE estimates are nearly identical.

Benjamin M. Craig, Ph.D.*
Assistant Member, Health Outcomes & Behavior Program,
Moffitt Cancer Center, Tampa, Florida, and
Courtesy Associate Professor, Department of Economics,
University of South Florida, Tampa, Florida, USA

*Corresponding Author, Moffitt Cancer Center, 12902 Magnolia Drive, MRC-CANCONT, Tampa, FL 33612-9416; Phone: (813) 745-6710; Fax: (813) 745-6525; benjamin.craig@moffitt.org

Keywords: QALY, EQ-5D, time trade-off, discrete choice

Word Count (1,883), Pages (7), References (2), Tables (0), Figures (4)

Weber’s QALY: TTO and DCE estimates are nearly identical

ABSTRACT

Background: According to Weber's law, the change in a stimulus that will be "just noticeable” is a constant ratio of the original stimulus (i.e., $dp=k(dS/S_0)$). After integration (i.e., $p = k \ln(S)+C$), this equation may be applied to the analysis of health valuation responses as the functional relationship between discrete choice experiment (DCE) probabilities and time trade-off (TTO) value estimates.

Methods: The likelihood of “noticing” was linked to perception (p) using a complementary log-log (CLL) function, 1-exp(-exp(p)). Under this simplifying assumption, Weber’s law may be expressed as $\ln(\text{Prob}(B>A)/0.5)= \alpha(\text{TTO}(B)-\text{TTO}(A))$, where α is a rescaling parameter. Agreement between the two sides of this conjecture (perception and stimulus) was visually and formally assessed (i.e., Lin’s coefficient of agreement and mean absolute difference).

Data: Using a standardized protocol developed by the EuroQol Group, rank and TTO responses were collected for 42 EQ-5D health states, optimal health, and dead in the United States (US) and United Kingdom (UK). For each pair of hypothesized states, we estimated probabilities, Prob(B>A), using rank responses and estimated the difference in state-specific TTO-based quality-adjusted life year (QALY) values, TTO(B)-TTO(A). The pairs were categorized as nested, non-nested, and anchored. Nested pairs are consecutive pairs of health states (e.g. 11111 and 11112), where dominance is logically defined by the EQ-5D descriptive system. Anchored pairs are single health states that are compared to optimal health or dead, states anchoring the QALY scale. The US valuation study contained 153 nested, 153 non-nested, and 84 anchored pairs. The UK valuation study contained 360 nested, 501 non-nested, and 84 anchored pairs.

Results: Based on the graphical evidence, Weber’s law is rejected for nested comparisons; however, the evidence supports Weber’s law for non-nested pairs. After rescaling perception ($\alpha = 6.18$ and 5.80, respectively), Lin’s coefficient of agreement is 0.873 and 0.858, respectively. Using a CLL model, Weber QALY estimates are nearly identical to the TTO estimates in the US and UK (Lin's coefficient of 0.935 and 0.957, respectively).

Conclusion: These promising results from US and UK suggest convergent validity between DCE and TTO valuation techniques. If replicated in other valuation studies, the simplicity of DCE may motivate its use over more complicated valuation techniques. This also suggests that addressing the TTO constant proportional assumption may marginally improve validity (a maximum increase in Lin's coefficient of 0.07).

Keywords: Rank, Quality of Life, EQ-5D, Time Trade-off

INTRODUCTION

Ernst Heinrich Weber (1795-1879) and Gustab Fechner (1801-1887) launched the field of psychophysics through their studies of human response to “just noticeable differences” in physical stimuli (Thurstone, 1927). Their experiment in weights, sound, and vision helped identify the smallest detectable difference between two sensory stimuli. This barely discernible difference is known as Weber’s constant. Above this threshold, Fechner conjectured that Weber’s law may apply:

> “In order that the intensity of a sensation may increase in arithmetical progression, the stimulus must increase in geometric progression”

The objective of this paper is to test Weber’s law in health valuation. Using a standardized protocol developed by the EuroQol Group, rank and time trade-off (TTO) responses were collected for 42 EQ-5D health states, optimal health, and dead in the United States (US) and United Kingdom (UK). We will test whether Weber’s law can serve as a link between the rank-based probability estimates (i.e., perception) and TTO-based value estimates (i.e., stimulus). If so, this evidence would demonstrate convergent validity of rank and TTO estimates, and provide a functional form to transform probabilities into quality-adjusted life years (QALYs) (i.e., Weber’s QALYs).

THEORY

Formally, Weber’s law may be expressed as $dp=k(dS/S_0)$, an equation known as Fechner’s law, where dp is the differential change in perception, dS is the increase in stimulus, and S is the instant stimulus. The integral of this equation (i.e., $p = k \ln(S)+C$) reveals a logarithmic relationship between perception and stimulus, where C is the constant of integration.

To apply Weber’s law to the analysis of health valuation responses, perception must be linked to the likelihood of “noticing.” This likelihood is a conditional probability and may be expressed as $1-P(B>A)/0.5$. The complementary log-log

(CLL) function, (1-exp(-exp(p)), was selected as the link function for its simplicity. Its complementary form simplifies the likelihood equation, P(B>A)/0.5= exp(-exp(p), and may be combined with the logarithmic relationship between perception and stimulus:

$$
\begin{aligned}
&\ln(\Pr(B > A)/0.5) = -\exp(p) \quad where \quad p = \ln(S) \\
&\ln(\Pr(B > A)/0.5) = -S \\
&\ln(\Pr(B > A)/0.5) = a\,(TTO(B) - TTO(A))
\end{aligned}
$$

where α is a rescaling factor, the constant factor, *k*, is one and the constant of integration, C, is zero. Under this specification, Fechner's law states that the log likelihood of not "noticing" is proportional to the difference in values.

Data

The Measurement and Value of Health (MVH) protocol, first developed and applied in the UK, has been described in detail elsewhere (Dolan, 1997; Gudex, 1994; Kind, et al., 1998) . In both country studies, health states are characterized using the EQ-5D system, which comprises a set of scores on five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) with three possible levels on each. A vector of these five scores may be used as shorthand in identifying specific health states. For instance, the abbreviation 21132 represents a health state with some problems in walking, no problems with self-care, no problems with performing usual activities, severe pain, and moderate anxiety.

The Measurement and Valuation of Health (MVH) Protocol

The original MVH-protocol describes a face-to-face interview with several sections. First, respondents are asked to describe their health using the EQ-5D system. Next, respondents place 15 or more cards, each describing a health state, in order by individual assessment. These 15 cards include the anchoring states of 11111 and "immediate death." Respondents are instructed to assume each health state persists for ten years, followed by death. Following the ranking exercise, the subjects are asked to place each card on the EQ-VAS.

In the TTO exercise that follows, respondents decide whether ten years in a health state is preferred to ‘immediate death.’ If so, a series of trade-offs are presented to determine the number of years, y_1, in optimal health (i.e., QALYs) that is equivalent to ten years in the state presented. If the respondent prefers `immediate death,’ the interviewer presents an alternative series of trade-offs to determine the number of QALYs, $y_2 < 10$, followed by years in the state presented, $(10-y_2)$, that together are equivalent to `immediate death.’ Because `immediate death’ has a value of zero on a QALY scale, $(10-y_2)$ years in the state presented is equivalent to $-y_2$ QALYs. Thus, the TTO exercise produces an amount of QALYs (either y_1 or $-y_2$) equal to years in the state presented (either 10 or $10-y_2$) depending on whether the state is regarded as better or worse than dead.

Exclusion of Respondents

Respondent data were excluded for a particular method based on any of the following three criteria: (1) only one or two states were valued (other than 11111, “immediate death”, and “unconscious”); (2) all states were given the same value; and (3) all states were valued worse than “immediate death.” In addition, respondents were excluded from the rank sample if they ranked death equivalent to optimal health. Among the 3,395 participants in the US MVH study, five did not provide TTO responses, 33 did not provide rank responses, 47 did not meet the TTO criteria, and 50 did not meet the rank criteria. Among the 4,048 respondents in the UK MVH study, 16 did not provide TTO, 16 did not provide rank responses, 47 did not meet the TTO criteria, and 46 did not meet the rank criteria. In total, these exclusions resulted in a less than 2% decrease in the analytical samples. No further logical consistency criteria were applied.

Estimation of TTO Values and Pairwise Probabilities

For each state, respondents reported the number of QALYs, y, equal to the years in the state presented, x. We stratified the sample by country and health state, and we estimated the TTO values using the coefficient approach proposed by Craig and Busschbach:

$$y_i = bx_i + e_i \quad \rightarrow \hat{b} = \sum_{i=1}^{N} x_i y_i \Big/ \sum_{i=1}^{N} x_i^{\,2}$$

The value estimator is the mean of xy over the mean of x^2, and compensates for measurement error in y. Furthermore, it may be justified under the assumption of an episodic random utility model.

To estimate pairwise probabilities, each rank response was exploded into a series of pairwise comparisons. The pair data were stratified by country and unique pair. In each pair, the first state is defined as the state more likely to be inferior to its counterpart; therefore, the estimated probability Prob(B>A) is always less than 0.5 by construction. Because of the conditionality assumption, the likelihood of not "noticing" is the estimated probability divided by 0.5.

Estimation of the Rescaling Parameter

Mapping from the probabilities to QALY differences or from QALY differences to probabilities requires a rescaling parameter (α). For each country, we estimated the mean of $\ln(\Pr(B > A)/0.5)$ and the mean of $(TTO(B) - TTO(A))$. To accommodate error in each side of the equation, the scaling parameter estimator is the ratio of these two means. This ratio of means approach is preferred to the mean of the ratio, $\ln(\Pr(B > A)/0.5)/(TTO(B) - TTO(A))$. Mean ratios become infinite as the denominator (i.e., difference in QALYs) approaches zero.

Weber's QALYs

If Weber's law holds, QALY values may be estimated using a complementary log-log (CLL) model.

$$\Pr(A > B) = 1 - 0.5\exp(Q_A - Q_B)$$

where Q_j is the coefficient of the indicator variable for state j and represents the product αQALY_j. This estimation requires comparisons to an anchor (e.g., dead) for identification, and ties may be incorporated into the estimator using the Efron approach.

After estimating the CLL model by likelihood maximization, the coefficient, Q_j, was rescaled to QALY values using the rescaling parameter estimate, α. The re-scaled coefficient estimates of the state-specific indicator variables (i.e., Weber's QALYs) were compared to the TTO estimates both visually and formally.

In complement to the graphs, we estimated Lin's coefficient of agreement to measure concordance with Weber's law and concordance between rescaled CCL predictions and TTO QALY estimates. This coefficient ranges from zero to one and measures distance from the 45 degree line. We also described agreement in terms of mean absolute difference on a QALY scale.

RESULTS

In the US study, a subset of respondents (N=20) ranked atypical pairs (120 pairs, each with a total of seven rank responses or less). The small sample sizes for these pairs prohibited the estimation of P(B>A), and these responses were removed from the analytical sample. The remaining US pairs had between 120 and 3,333 rank responses. The US analytical study contained 153 nested, 153 non-nested, and 84 anchor pairs. In the UK sample, the sample size ranged from 375 and 3,965 rank responses per pair, and the analytical sample contained 360 nested, 501 non-nested, and 84 anchor pairs.

Figure 1 illustrates the relationship between the probability of a health state (H) dominating an anchor state (i.e., P(H > anchor state)) and the estimated TTO value of the health state. Based on US and UK estimates, there is no association between P(H > optimal health) and the estimated TTO value. In optimal health, the highest probability estimate is 1.5% in the US and 2.2% in the UK, which suggests that discrete choice experiment (DCE) probabilities with anchor states provide no information on health state value. The association between P(H > dead) and the estimated TTO value appear similar in the US and UK studies; however, the states with TTO estimates near zero are more likely to dominate death, a pattern that violates Weber's law. This violation is an ordinal contradiction: health states below dead in TTO value are considered better than dead in rank.

Figures 2 illustrates the relationship between ln((B>A)/0.5) and TTO(B)-TTO(A) among the nested pairs in the US and UK studies. The evidence suggests a linear relationship, but rejects proportionality (i.e., the pattern does not fit through the origin). The nesting implies a difference in stimulus; however, it also catalyzes the perception of this stimulus, leading to a downward shift in the log probability. In other words, nested states appear more dissimilar in rank than in TTO value.

Figure 3 illustrates the relationship among the non-nested pairs in the US and UK studies. The visual evidence from the US and UK suggest proportionality, which is complemented by the formal results. After rescaling (α = 6.18 and 5.80, respectively), Lin's coefficient of agreement is 0.873 and 0.858, respectively, and the absolute mean difference between QALY differences is 0.065 and 0.081, respectively, which suggests that DCE and TTO QALY estimates are nearly identical.

Figure 4 shows the relationship between Weber's QALY predictions and TTO QALY predictions for the 42 health states based on the dead and non-nested pair responses in the US and UK studies. Lin's coefficient of agreement is 0.935 and 0.957, respectively, and the absolute mean difference is 0.073 and 0.062, respectively. In addition to showing the near equivalence of the TTO and DCE estimates, the pattern demonstrates the feasibility of the CLL estimator to produce state-specific QALY values with strong convergent validity using only dead and non-nested pair comparisons responses.

Summary and Conclusion

Nearly identical patterns are found in the US and UK estimates. Weber's law under the CLL assumption appears to hold for non-nested comparisons. For dead comparisons, Weber's law appears to be violated, but this lack of convergent validity may be attributed to issues with the TTO valuation technique. Weber's QALY estimates are nearly identical to TTO estimates, which support the use of DCE-based QALY estimation. Although promising, the estimation of QALYs using DCE requires the assumption of a rescaling parameter, α, which remains controversial.

Nested comparisons using DCE contain deterministic and random utility components. Once the deterministic component is controlled, further work may allow the incorporation of these responses into the QALY estimation. However, DCE comparisons with optimal health are largely deterministic (likely attributable to logical dominance). These comparisons provide little information on random utility (i.e., non-optimal gap).

REFERENCES

1. Craig BM, Busschbach JJ. The episodic random utility model unifies time trade-off and discrete choice approaches in health state valuation. Population Health Metrics 2009; 7; 3.
2. Dolan P. Modeling valuations for EuroQol health states. Medical Care 1997; 35; 1095-1108.
3. Gudex C. Time Trade-Off User Manual: Props and Self-Completion Methods. Report of the Centre for Health Economics. United Kingdom: York, University of York; 1994.
4. Kind P, Dolan P, Gudex C, Williams A. Variations in population health status: results from a United Kingdom national questionnaire survey. British Medical Journal 1998; 316; 736-741.
5. Thurstone LL. Three psychophysical laws. Psychological Review 1927; 34; 424-432.

Figure 1. Probability of Dominating the Anchor State and the TTO Value for 42 EQ-5D States

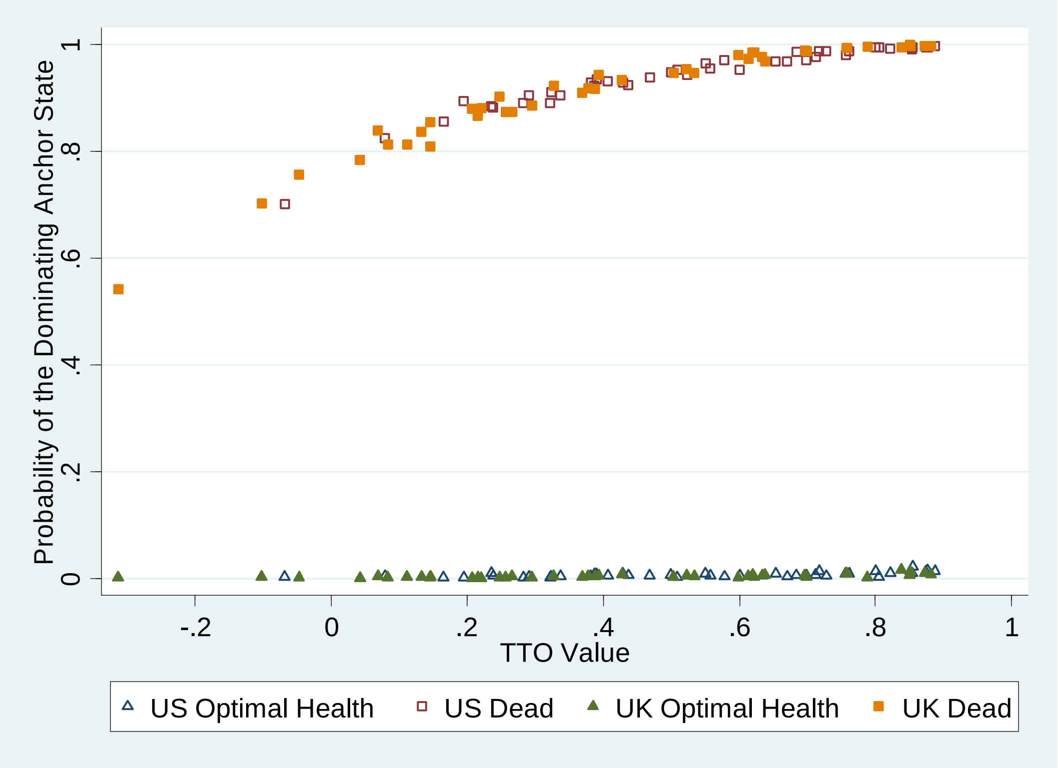

Figure 2. Perception and Stimulus among Nested EQ-5D Pairs

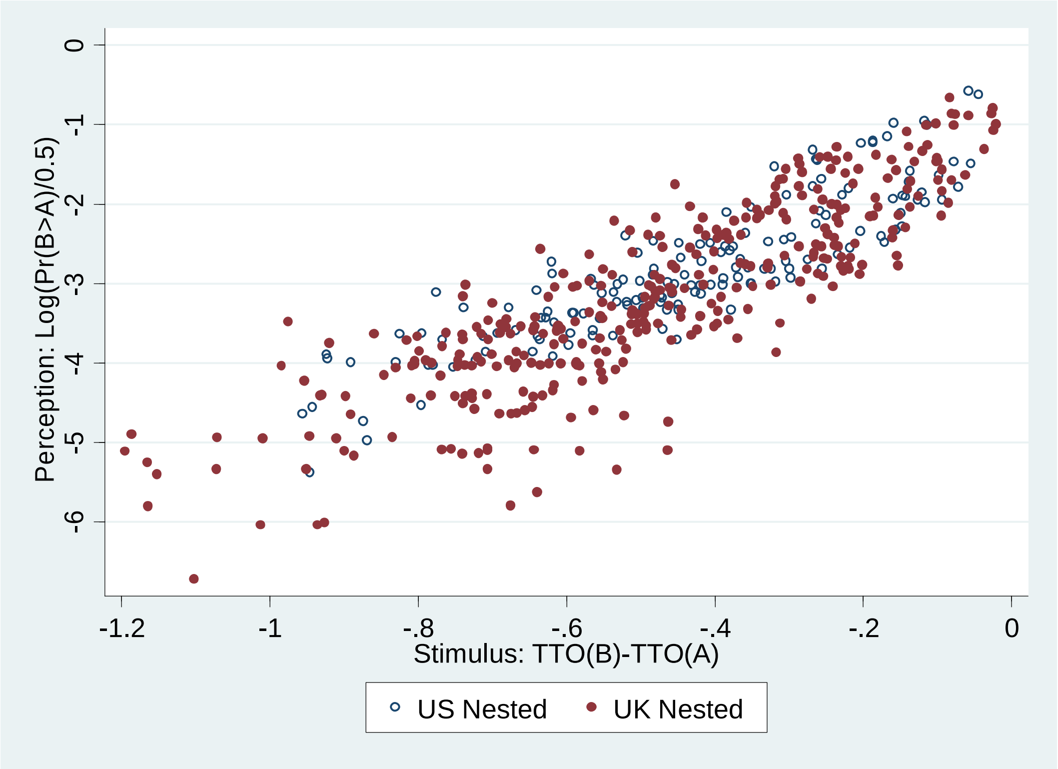

Figure 3. Perception and Stimulus among Non-Nested EQ-5D Pairs

Figure 4. Weber's QALY and TTO Estimates for 42 EQ-5D States