Non-monotonicity in the episodic random utility model

Nicolas A Menzies[1] and **Joshua A Salomon**[1,2]

[1]Center for Health Decision Sciences, Harvard University, Boston, MA, USA

[2]Department of Global Health and Population, Harvard School of Public Health, Boston, MA, USA

Abstract

The time trade-off is widely used in population-based surveys to estimate health-state valuations. Typically, respondents may characterize states as being better than or worse than dead. However, worse-than-dead responses can produce strongly negative mean values, so various analytic transformations of these responses have been suggested. The episodic random utility model (eRUM), operationalized using a linear regression estimator, was proposed as an alternative to these transformations, in part because of its theoretical appeal. We analyzed the eRUM estimator's mathematical properties and found that it violates monotonicity under certain patterns of survey responses, such that improvement in some individual valuations would imply a lower overall valuation for a given health state. Consequently, it is possible that orderings of alternative strategies based on eRUM valuations could lead a decision-maker to choose a strictly dominated strategy. Re-analyzing data from a large population-based EQ-5D valuation survey in the United Kingdom, we found 27% of all time trade-off responses (63% of all worse-than-dead responses) met the conditions for violation of monotonicity, and 74% of all respondents had at least one such response. These results present some challenge to the use of the eRUM estimator in generating health-state valuations for population health measurement and economic evaluation.

Keywords

Quality-Adjusted Life Years; Quality of Life; Statistical Model; Utility Theory

1. Introduction

Health-state valuations represent important inputs to cost-effectiveness analysis and related research (Gold *et al.,* 1996; Drummond and McGuire, 2001). Valuation approaches vary, but guidelines generally focus on elicitation methods that force respondents to make explicit choices, such as the time trade-off (TTO) and standard gamble techniques (Gold *et al.,* 1996; NICE, 2008). The TTO is used widely in deriving valuations for generic health-state descriptions through sample surveys in the general public (Dolan *et al.,* 1996; Tsuchiya *et al.,* 2002; Shaw *et al.,* 2005). A common feature in many TTO studies is that respondents may characterize states as being either better than dead (BTD) or worse than dead (WTD) (Torrance, 1982; Patrick *et al.,* 1994; Tilling *et al.,* 2010). Allowing for WTD valuations presents a range of conceptual, methodological and empirical challenges that have been met with varying proposed solutions, but debate persists about the most appropriate approach (Tilling *et al.,* 2010; Robinson and Spencer, 2006; Lamers, 2007). In this paper, we present a critical evaluation of the episodic random utility framework, proposed recently as a new

Corresponding Author: Nicolas Menzies, Center for Health Decision Sciences, Harvard University, Boston, MA, USA. (+1) 404-217-1076 (phone), 617-432-0190 (fax), nmenzies@fas.harvard.edu.

paradigm for understanding and analyzing health-state valuations, with particular relevance for accommodating WTD TTO responses (Craig and Busschbach, 2009; Craig and Busschbach, 2010).

2. Interpretation of Time Trade-Off Responses

One of the most widely used protocols for eliciting TTO responses was introduced in the landmark Measurement and Valuation of Health (MVH) study in the United Kingdom (MVH Group, 1994), and provides a concrete focal example for understanding the challenge of WTD responses. In the MVH protocol, valuation of a health state using the TTO begins with a question to determine whether the state is regarded as BTD or WTD. For a state deemed BTD, respondents are asked a series of binary questions to determine how many years of life lived in perfect health followed by immediate death would be equivalent to 10 years of life lived in the health state followed by immediate death. For a state deemed WTD, respondents are asked questions to determine the point of equivalence between immediate death and the alternative of some number of years spent in the health state followed by the remaining balance of a 10-year life expectancy in perfect health. For both BTD and WTD valuations, choices are offered in in ¼ year increments (MVH Group, 1994).

Precedents for the MVH protocol are found in the original formulation of the TTO (Torrance *et al.*, 1972) and its subsequent elaboration to accommodate WTD valuations (Torrance, 1982). These seminal works also describe how individual TTO responses can be translated into valuations on an interval scale. Aggregate values for groups are typically calculated as the arithmetic mean of these individual valuations. Applying the original logic from Torrance *et al.* (1972) to the 10-year formulation of the TTO in the MVH, individual valuations for BTD states are calculated as $y/10$, where y is the number of years spent in perfect health equivalent to 10 years in the health state. The resulting BTD valuations are thus located within the interval (0,1). For WTD states, the formula suggested by Torrance (1982) implies a valuation of $y/(y-10)$, which allows WTD values to fall in the interval [− , 0). In practice the lowest possible valuation using the MVH protocol would be −39 (due to the ¼ year increments), which results from a WTD response in which a person is indifferent between immediate death and an alternative of 0.25 years in the health state followed by 9.75 years in perfect health. Calculating WTD observations as $y/(y-10)$, while consistent with the original interpretation of the equivalence statements elicited through the TTO, has led to unsatisfactory results when applied to survey data, with a large number of health states determined to have negative average valuations, challenging face validity (Dolan *et al.,*1996).

A common response to this problem has been to transform WTD observations so that they fall within the range [−1,0). The earliest proposed approach (Torrance, 1982; Torrance *et al.*, 1982), which was later adopted by Shaw *et al.* (2005), was to divide all WTD observations by a constant that inflates the minimum valuation to −1 (and all other WTD valuations by the same factor). Applied alongside the MVH protocol, this transformation implies WTD valuations equal to $\left(1/39\right)\left(y/(y-10)\right)$. Another frequently used transformation was introduced by Patrick *et al.* (1994) and applied in the United Kingdom MVH study and a number of subsequent studies. Describing this latter transformation, Dolan *et al.* (1996) reasoned that respondents were misinterpreting the WTD response scale as an interval rather than a ratio scale, citing early psychometric research (Eyman, 1967), and calculated a modified individual valuation for WTD observations, as $-y/10$. In all cases the calculation

for BTD valuations is unchanged, and these estimators produce final valuations that are bounded on the scale [−1,1], with minor differences between estimators due to the particular transformation used (Lamers, 2007). These transformations are pragmatic approaches for developing an estimator based on the mean of a set of individual responses, which respects the orderings in these responses (more favorable responses imply higher valuations), while also generating plausible results. However, they have been criticized for lacking a strong theoretical basis and for their reliance on ratio statistics (Robinson and Spencer, 2006; Lamers, 2007; Craig and Busschbach, 2009).

Craig and Busschbach (2009) suggest an alternate approach to calculating credible valuations directly from the unadjusted observations, by reinterpreting the source of stochastic variability in individual survey responses. This alternative is developed formally as the episodic random utility model (eRUM), whereby the reported utility (U) of an individual (i) for a health state (j) over time (t) involves a systematic component (μt) that depends on the amount of time spent in the state, and a stochastic component ( ) that does not. The eRUM differs from the traditional instant random utility model (iRUM) in its conceptualization of the error term:

$$U_{ij}(t) = \begin{cases} \mu_j t + \varepsilon_{ij} & eRUM \\ (\mu_j + \varepsilon_{ij})t = \mu_j t + \varepsilon_{ij} t & iRUM \end{cases} \quad (\text{Craig and Busschbach}, 2009)$$

Considering the implications of the eRUM framework for deriving health-state valuations, Craig and Busschbach note that eRUM is consistent with a linear regression estimator, in which the number of years of perfect health (multiplied by −1 for WTD responses) is regressed against the number of years in a particular health state, with no constant term. Under eRUM the regression coefficient, rather than the mean of individual valuations, is the estimator for the valuation of the given health state. Craig and Busschbach have recently reiterated the regression-based operationalization as a general approach for analysis of health-state valuations, without explicit reference to the underlying eRUM model (Craig and Busschbach, 2010).

3. Non-Monotonicity of the Craig & Busschbach Estimator

Non-monotonicity is an undesirable property for an aggregate measure of social preference (Arrow, 1950). In the context of health-state valuations, if some respondents improve their valuation of a particular state, it is undesirable that the aggregate measure of value declines as a result. We hypothesized that the estimator derived from the eRUM could violate monotonicity. This hypothesis was based on the observation that, for health states that are generally valued positively, the influence of WTD scores on the slope of the regression line —constrained to pass through the origin—must decrease for increasingly negative valuations. At the extreme, if an individual were allowed to value a health state as infinitely bad (i.e. with a raw score of −10 implying an individual valuation of −), this observation would exert no influence on the regression line at all. Analogous to the physical concept of torque, the infinitely negative valuation produces no downward rotation of the regression line because the force is applied at zero horizontal distance to the pivot point (at the origin).

To test our hypothesis, we sought to derive the mathematical conditions that would characterize violations of monotonicity. Following Craig and Busschbach, we began by describing the eRUM estimator (μ) for the valuation of a particular health state in relation to the set of BTD and WTD observations, where x is the number of years spent in the health state and y is the number of years spent in perfect health (recorded as a negative value for WTD observations).

$$\widehat{\mu}=\frac{\sum\limits_{i} x_i y_i}{\sum\limits_{i} x_i^2}=\frac{\sum\limits_{BTD} 10y+\sum\limits_{WTD} (y+10)y}{\sum\limits_{BTD} 10^2+\sum\limits_{WTD} (y+10)^2}$$

The monotonicity condition may be understood in terms of the direction of change in μ caused by a change in a single observation, so we split off one WTD observation, *p*:

$$\widehat{\mu}=\frac{\sum\limits_{BTD} 10y+\sum\limits_{WTD-p} (y+10)y+y_p(y_p+10)}{\sum\limits_{BTD} 10^2+\sum\limits_{WTD-p} (y+10)^2+(y_p+10)^2}=\frac{\left[\sum\limits_{BTD} 10y+\sum\limits_{WTD-p} (y+10)y\right]+y_p^2+10y_p}{\left[\sum\limits_{BTD} 10^2+\sum\limits_{WTD-p} (y+10)^2\right]+y_p^2+20y_p+100}=\frac{f(y_p)}{g(y_p)}$$

Next, we took the partial derivative of this expression with respect to y_p,

$$\frac{\partial\widehat{\mu}}{\partial y_p}=\frac{f'(y_p)g(y_p)-f(y_p)g'(y_p)}{g(y_p)^2}=\frac{f'(y_p)-\widehat{\mu}\,g'(y_p)}{g(y_p)}$$

Monotonicity requires that the regression estimator μ must increase in y_p; that is,

$$\frac{\partial\widehat{\mu}}{\partial y_p}>0,\text{ or }\frac{f'(y_p)-\widehat{\mu}g'(y_p)}{g(y_p)}>0.$$

Noting that $g(y_p)$ is the denominator of the original regression equation, computed as the sum of squared values and therefore always positive, we restated the monotonicity condition based only on the numerator, which led to a simple statement of the monotonicity condition.

$$f'(y_p)-\widehat{\mu}\,g'(y_p)>0$$
$$\Rightarrow \frac{f'(y_p)}{g'(y_p)}>\widehat{\mu}$$
$$\Rightarrow \frac{y_p+5}{y_p+10}>\widehat{\mu}$$

The region of the joint distribution of y_p and μ where this monotonicity condition fails is shown in Figure 1. Of note, for values of μ greater than 0.5, the improvement of *any* WTD observation will lower the value of the eRUM estimator, violating monotonicity.

One implication of this monotonicity violation is that a state could dominate another state by first-order stochastic dominance (Hardar and Russell, 1969) yet have a lower eRUM estimator, as shown by the stylized 4-observation example in Table I. In this example, an intervention that moved patients from Health State A to Health State B would, *ceteris parabis*, be judged a social welfare improvement according to the eRUM estimate, even though this makes all individuals worse-off and violates Pareto-efficiency.

Expanding on this theoretical concern, we considered the practical implications of monotonicity violations using data from the United Kingdom MVH study (Dolan *et al.,* 1996). The MVH study collected TTO valuations for 43 health states described using the EQ-5D. The EQ-5D system includes five domains (mobility, self-care, usual activities, pain/

discomfort, and anxiety/depression), each with three possible levels. States are described in shorthand using five-digit codes reflecting levels on each domain.

We estimated the percentage of observations in the MVH study that meet the conditions for non-monotonicity, and results are shown in Table II. Overall, 27% of all individual observations (63% of all WTD observations) fall into the region affected by non-monotonicity. For individual health states, the percentages of observations in the affected region range from 2% to 54% of all observations (58% to 100% of all WTD observations). Figure 2 shows the distribution of observations for six example health states, with the observations in the non-monotonic region shaded red.

Respondents in the MVH study were typically asked 13 TTO questions, and 74% of all respondents provided at least one response falling into the affected region; 44% provided 4 or more responses falling into the affected region. Comparing the distribution of raw scores across health states, we identified no instances where a dominated health state received a higher eRUM estimate.

4. Conclusion

This analysis investigated properties of the episodic random utility estimator (Craig and Busschbach, 2009; Craig and Busschbach, 2010) used to model survey data on health-state valuations. We found that the eRUM estimator can violate monotonicity under certain patterns of survey responses, and this potential problem pertains to a substantial proportion of all observations in a large population-based survey of health-state valuations in the United Kingdom. This problem is clearly not restricted to extreme values or outliers—for one-quarter of all observations, an incremental improvement in the individual valuation of a particular health state would result in a worsening of the aggregate valuation for that state derived using the eRUM approach. Moreover, almost three-quarters of all respondents provided at least one such response.

The eRUM approach, like the set of strategies based on transformations of WTD observations (Torrance, 1982; Dolan *et al.,* 1996; Shaw *et al.,* 2005), is proposed as a possible alternative to the original analytic strategy designed alongside the TTO (Torrance *et al.,* 1972). Reacting to the implausibility of strongly-negative valuations produced by this original approach, both the eRUM and the transformation-based alternatives invoke strong assumptions to justify their particular analytic choices. Both produce the desired effect of eliminating implausibly large negative valuations, and the rankings of health states resulting from the two approaches are quite similar. This raises a reasonable question—is there really a problem with the eRUM estimator if it produces plausible results in practice? While the close concordance of health-state valuations produced by the different estimators renders the choice between them less consequential, an ideal estimator should be justifiable theoretically as well as empirically. Indeed, the eRUM is explicitly touted by its proponents as being theoretically superior to other estimators (Craig and Busschbach, 2009; Craig and Busschbach, 2010). The issues revealed by our analysis present some challenge to this position.

For the analyst weighing eRUM against the transformation approach, neither strategy emerges as clearly dominant. Both have shortcomings, and these shortcomings relate to distinct types of considerations. Arguments against the selective transformation of WTD observations have been rehearsed extensively in the literature (Lamers, 2007; Craig and Busschbach, 2009; Craig and Oppe, 2010; Shaw *et al.,* 2010a). The regression-based estimator used in eRUM has been introduced more recently, and has not yet been subjected to the same scrutiny. In this paper, we identify an important failing of the eRUM approach,

and based on empirical re-analysis of a prominent national valuation study find that this problem is highly relevant to observed, real-world response patterns. Thus, a choice between these two imperfect alternatives must ultimately depend on the relative importance that an analyst attaches to concerns about internal consistency or arbitrariness in transformation approaches vs. concerns about monotonicity violations in the eRUM estimator. Furthermore, it is worth noting that neither alternative fully reconciles known features of TTO data, such as the gap in valuations in the proximity of the dead state (Stalmeier, 2005), the greater variability of WTD observations (Dolan and Roberts, 2002), and the poor correlation between mean negative scores and overall valuations (Busschbach *et al.,* 2003).

Additional alternatives for analyzing health-state valuation data continue to emerge, such as a strategy focusing on median rather than mean valuations (Lamers, 2007; Li and Fu, 2009; Shaw *et al.,* 2010a), and an estimator based on directional statistics (Craig and Oppe, 2010) which—like the eRUM estimator—avoids the use of ratio statistics. At present, however, consensus on an ideal approach remains elusive (Shaw *et al.,* 2010b). Under these circumstances, research on revised TTO question formats (e.g., Robinson and Spencer, 2006; Buckingham and Spencer, 2006; Devlin *et al.,* 2010) and renewed consideration of ranking and other ordinal elicitation methods (e.g., Salomon, 2003; McCabe *et al.,* 2006; Krabbe *et al.,* 2007; Craig *et al.,* 2009) merit continued attention in seeking to avoid or ameliorate problems with WTD health-state valuations associated with the TTO.

Acknowledgments

References

Arrow KJ. A difficulty in the concept of social welfare. Journal of Political Economy. 1950; 58(4): 328–346.

Buckingham K, Devlin N. A theoretical framework for TTO valuations of health. Health Economics. 2006; 15(10):1149–1154. [PubMed: 16786547]

Busschbach, J.; Weijnen, T.; Nieuwenhuizen, M.; Oppe, S.; Badia, X.; Dolan, P., et al. A comparison of EQ-5D time trade-off values obtained in Germany, the United Kingdom and Spain. In: Brooks, R.; Rabin, R.; de Charro, F., editors. The measurement and valuation of health status using EQ-5D: a European perspective. Dordrecht: Kluwer Academic Publishers; 2003.

Craig BM, Busschbach JJV. The episodic random utility model unifies time trade-off and discrete choice approaches in health state valuation. Population Health Metrics. 2009; 7:3. [PubMed: 19144115]

Craig BM, Busschbach JJV. Toward a more universal approach in health valuation. Health Economics. 2010

Craig BM, Busschbach JJV, Salomon JA. Keep it simple: ranking health states yields values similar to cardinal measurement approaches. Journal of Clinical Epidemiology. 2009; 62(3):296–305. [PubMed: 18945585]

Craig BM, Oppe M. From a different angle: a novel approach to health valuation. Social Science & Medicine. 2010; 70:169–174. [PubMed: 19880235]

Dolan P, Gudex C, Kind P, Williams A. The time trade-off method: results from a general population study. Health Economics. 1996; 5:141–154. [PubMed: 8733106]

Devlin NJ, Tsuchiya A, Buckingham K, Tilling C. A uniform time trade off method for states better and worse than dead: feasibility study of the 'lead time' approach. Health Economics. 2010 Published Online.

Dolan P, Roberts J. To what extent can we explain time trade-off values from other information about respondents. Social Science & Medicine. 2002; 54:919–929. [PubMed: 11996025]

Drummond, M.; McGuire, A. Economic Evaluation in Healthcare: Merging Theory with Practice. Oxford, UK: Oxford University Press; 2001.

Eyman RK. The effect of sophistication on ratio and discriminative scales. American Journal of Psychology. 1967; 80:520–540. [PubMed: 6081391]

Gold, MR.; Siegel, JE.; Russell, LB.; Weinstein, MC., editors. Cost-Effectiveness in Health and Medicine. New York, USA: Oxford University Press; 1996.

Hardar J, Russell WR. Rules for ordering uncertain prospects. American Economic Review. 1969; 59(1):25–34.

Krabbe PFM, Salomon JA, Murray CJL. Quantification of health states with rank-based nonmetric multidimensional scaling. Medical Decision Making. 2007; 27(4):395–405. [PubMed: 17761959]

Lamers LM. The transformation of utilities for health states worse than death: consequences for the estimation of EQ-5D value sets. Medical Care. 2007; 45:238–244. [PubMed: 17304081]

Li L, Fu AZ. Some methodological issues with the analysis of preference-based EQ-5D index score. Health Services and Outcomes Research Methodology. 2009; 9(3):162–176.

McCabe C, Brazier J, Gilks P, Tsuchiya A, Roberts J, O'Hagan A, Stevens K. Using rank data to estimate health state utility models. Journal of Health Economics. 2006; 25(3):418–431. [PubMed: 16499981]

MVH Group. Time Trade-Off User Manual: Props and Self-Completion Methods. York, UK: University of York Center for Health Economics; 1994.

NICE. Guide to the Methods of Technology Appraisal. London, UK: National Institute for Health and Clinical Excellence; 2008.

Patrick DL, Starks HE, Cain KC, Uhlmann RF, Pearlman RA. Measuring preferences for health states worse than death. Medical Decision Making. 1994; 14:9–18. [PubMed: 8152361]

Robinson A, Spencer A. Exploring challenges to TTO utilities: valuing states worse than dead. Health Economics. 2006; 15(4):393–402. [PubMed: 16389652]

Salomon JA. Reconsidering the use of rankings in the valuation of health states: a model for estimating cardinal values from ordinal data. Population Health Metrics. 2003; 1:12. [PubMed: 14687419]

Shaw JW, Johnson JA, Coons SJ. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. Medical Care. 2005; 43(3):203–220. [PubMed: 15725977]

Shaw JW, Picard AS, Yu S, Chen S, Iannacchione VG, Johnson JA, Coons SJ. A median model for predicting United States population-based EQ-5D health state preferences. Value in Health. 2010a; 13(2):278–288. [PubMed: 19961566]

Shaw JW, Pickard AS, Yager RR. Application of directional statistics to health state valuation: a commentary on Craig and Oppe. Social Science & Medicine. 2010b; 71:429–430. [PubMed: 20553854]

Stalmeier PFM, Busschbach JJV, Lamers LM, Krabbe PFM. The gap effect: discontinuities of preferences around dead. Health Economics. 2005; 14:679–685. [PubMed: 15744750]

Tilling C, Devlin N, Tsuchiya A, Buckingham K. Protocols for time tradeoff valuations of health states worse than dead: a literature review. Medical Decision Making. 2010 Published Online.

Torrance, GW. Multiattribute utility theory as a method of measuring social preferences for health states in long-term care. In: Kane, RL.; Kane, RA., editors. Values and Long-Term Care. Lexington, MA: Lexington Books Division of D. C. Health; 1982. p. 127-156.

Torrance GW, Boyle MH, Horwood SP. Application of multi-attribute utility theory to measure social preferences for health states. Operations Research. 1982; 30(6):1043–1069. [PubMed: 10259643]

Torrance GW, Sackett D, Thomas W. A utility maximization model for evaluation of health care programmes. Health Services Research. 1972; 7(2):118–133. [PubMed: 5044699]

Tsuchiya A, Ikeda S, Ikekami N, Nishimura S, Sakai I, Fukuda T, Hamashima C, Hisashige A, Tamura M. Estimating an EQ-5D population value set: the case of Japan. Health Economics. 2002; 11:341–345. [PubMed: 12007165]

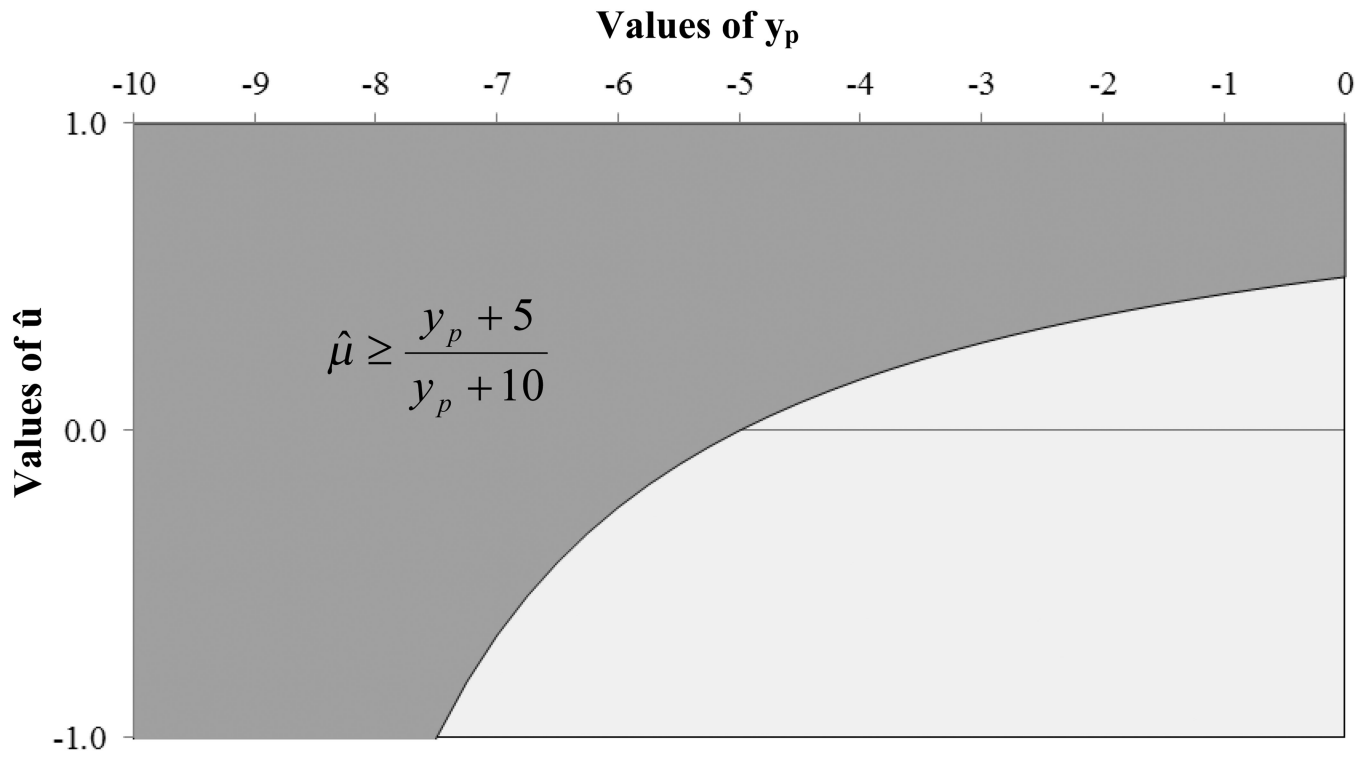## Values of y_p

Figure 1. Distribution of values of μ and y_p that violate monotonicity

μ = eRUM estimator for a given health state

y_p = years lived in perfect heath from a single survey response for the given health state (multiplied by -1 for states deemed worse than dead).

Dark grey area = joint distribution of y_p and μ associated with monotonicity violations.

Distribution of Responses for Health State 21111

2.0% of all responses non-monotonic

Distribution of Responses for Health State 22121

8.2% of all responses non-monotonic

Distribution of Responses for Health State 11113

26.0% of all responses non-monotonic

Distribution of Responses for Health State 21232

36.2% of all responses non-monotonic

Distribution of Responses for Health State 22233

39.7% of all responses non-monotonic

Distribution of Responses for Health State 33333
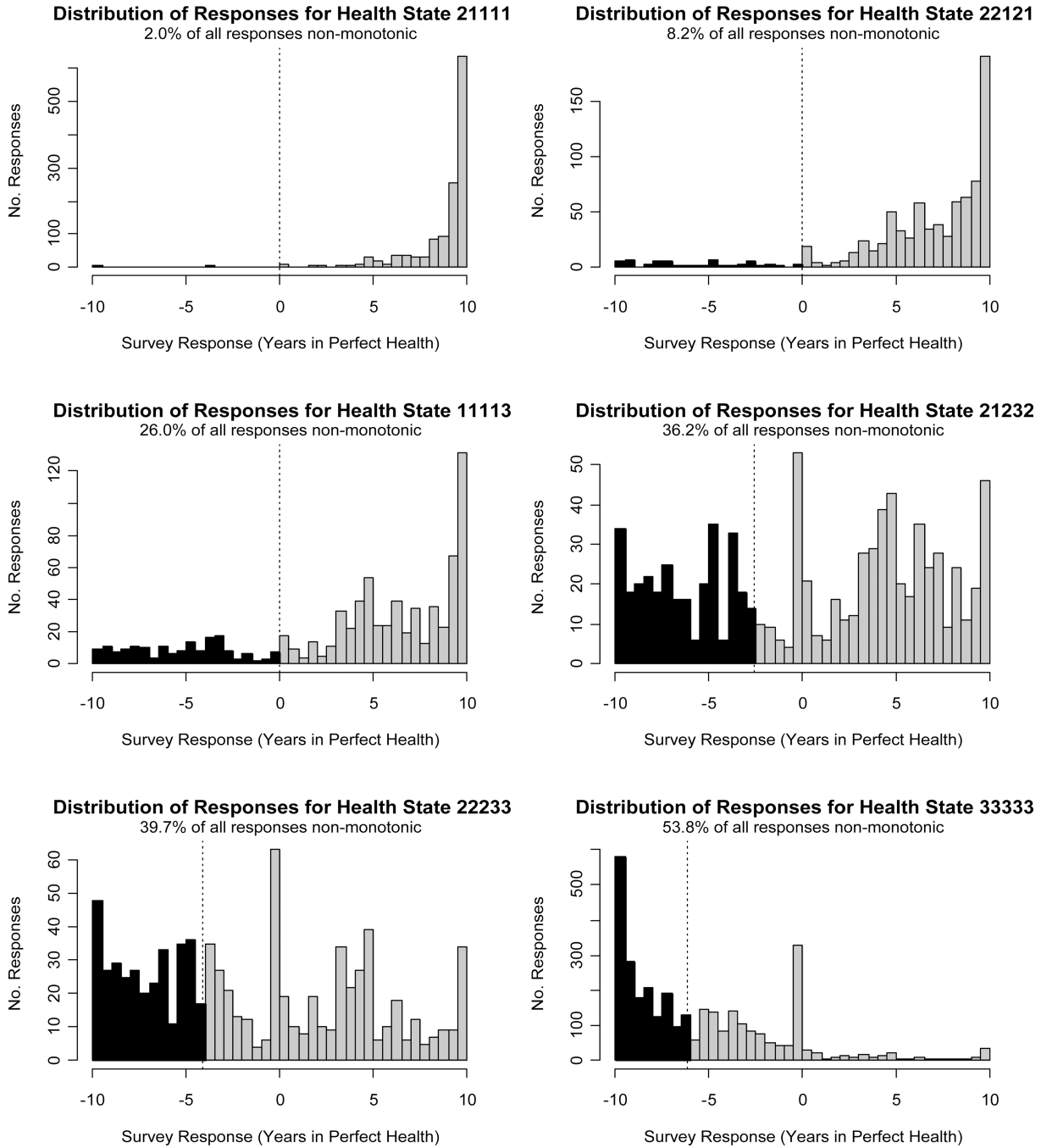
53.8% of all responses non-monotonic

Figure 2. Distribution of survey responses for six example health states in the United Kingdom Measurement and Valuation of Health study. Responses meeting the conditions for monotonicity violations are shaded in black

Table I

Example of possible violation of first-order stochastic dominance with eRUM

Observation	Health State A (Dominant)		Health State B (Dominated)	
	x-values	y-values	x-values	y-values
1	10	10	10	9.75
2	10	9	10	8.75
3	10	8	10	7.75
4	9	−1	6	−4
eRUM estimator:	0.685		0.709	

x-values = no. years spent in the health state y-values = no. years spent in perfect health

Table II

Percentage of observations meeting the conditions for monotonicity violations, for EQ-5D health states in the United Kingdom Measurement and Valuation of Health study.

EQ-5D health state	eRUM estimate	No. observations	Percent of observations WTD	Percent of observations in non-monotonic region	Percent of WTD observations in non-monotonic region
21111	0.88	1,319	2.0%	2.0%	100%
11211	0.87	1,347	1.7%	1.7%	100%
11121	0.85	1,324	2.6%	2.6%	100%
12111	0.85	1,322	3.6%	3.6%	100%
11112	0.84	1,318	3.5%	3.5%	100%
12211	0.79	834	4.4%	4.4%	100%
11122	0.76	827	6.3%	6.3%	100%
12121	0.76	835	4.2%	4.2%	100%
22121	0.70	837	8.2%	8.2%	100%
22112	0.70	847	6.5%	6.5%	100%
11312	0.64	832	14.7%	14.7%	100%
21222	0.63	828	12.7%	12.7%	100%
12222	0.62	837	15.4%	15.4%	100%
22122	0.62	820	14.1%	14.1%	100%
21312	0.61	816	15.7%	15.7%	100%
22222	0.60	842	16.5%	16.5%	100%
11113	0.53	832	26.0%	26.0%	100%
13212	0.52	826	22.3%	22.3%	100%
13311	0.50	816	27.1%	27.1%	100%
11131	0.43	819	37.5%	31.5%	84.0%
12223	0.40	836	37.2%	28.1%	75.6%
23321	0.39	825	40.2%	32.7%	81.3%
21323	0.38	826	39.2%	31.5%	80.2%
32211	0.37	842	40.1%	31.9%	79.6%
21232	0.33	831	46.5%	36.2%	78.0%

EQ-5D health state	eRUM estimate	No. observations	Percent of observations WTD	Percent of observations in non-monotonic region	Percent of WTD observations in non-monotonic region
22323	0.30	819	47.9%	35.0%	73.2%
22331	0.27	820	52.0%	38.9%	74.9%
33212	0.26	833	52.8%	37.6%	71.1%
11133	0.25	837	55.0%	39.9%	72.6%
21133	0.22	832	57.2%	38.2%	66.8%
23313	0.22	836	55.3%	38.5%	69.7%
23232	0.21	831	59.0%	41.0%	69.6%
22233	0.15	834	62.5%	39.7%	63.5%
33321	0.15	834	63.7%	37.3%	58.6%
32313	0.14	842	64.4%	38.8%	60.3%
32223	0.11	831	65.9%	39.0%	59.1%
32232	0.09	825	69.7%	43.4%	62.3%
13332	0.07	824	69.1%	42.4%	61.3%
32331	0.05	832	73.1%	41.3%	56.6%
33232	−0.04	836	77.8%	42.1%	54.2%
33323	−0.10	844	82.0%	43.6%	53.2%
Unconscious	−0.21	3,327	96.2%	35.2%	36.6%
33333	−0.30	3,320	93.5%	53.8%	57.5%