



## NIH PUBLIC ACCESS

## Author Manuscript

*Soc Sci Med.* Author manuscript; available in PMC 2011 January 1.

Published in final edited form as:

*Soc Sci Med.* 2010 January ; 70(2): 169–174. doi:10.1016/j.socscimed.2009.10.009.

## From a Different Angle: a novel approach to health valuation

**Benjamin M Craig** and

Moffitt Cancer Center Tampa, FL UNITED STATES [Proxy] Benjamin.Craig@moffitt.org

**Mark Oppe, MSc**

### Abstract

The value of a health state is typically described relative to the value of an optimal state, specifically as a ratio ranging from unity (equal to optimal health) to negative infinity. Incorporating potentially infinite values is a challenging issue in the econometrics of health valuation.

In this paper, we apply a directional statistics approach based on the assumption of wavering preference. Unlike ratio statistics, directional statistics are based on polar coordinates (angle, radius). The range of angles is bounded between 45 degrees (unity) and negative 90 degrees (i.e., negative infinity); therefore, mean angles are well behaved and negate the impetus behind arbitrary data manipulations. Using time trade-off (TTO) responses from the seminal Measurement and Valuation of Health study, we estimate 243 EQ-5D health state values by minimizing circular variance with and without radial weights.

For states with published values greater than zero (i.e., better-than-death), the radially weighted estimates are nearly identical to the published values (Mean Absolute Difference 0.07; Lin's rho 0.94). For worse-than-death states, the estimates are substantially lower than the published values (Mean Absolute Difference 0.186; Lin's rho 0.576). For the worst EQ-5D state (33333), the published value is -0.59 and the directional estimate is -1.11.

By taking a directional statistics approach, we circumvent problems inherent to ratio statistics and the systematic bias introduced by arbitrary data manipulations. The predictions suggest that published estimates overvalue severe states. This paper examines TTO responses; however, it may be extended to all forms of health valuation.

### Keywords

Rank; Quality of Life; EQ-5D; Time Trade-off

## INTRODUCTION

Can you imagine a health state where you would rather die than be in that state? If not, might another? Respondents with infinitely negative values for specific health states are commonly encountered in health valuation studies. When accumulating preferences across individuals, such extreme values negate the validity of summary statistics, such as population means and variances. A single infinite value causes the statistic to become infinite itself, losing all information about the other values within the population.

© 2009 Elsevier Ltd. All rights reserved.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Even if such an extreme value is impossible, the potential of an extreme response (i.e., stated value) hinders survey research. When asked in a survey, respondents may state that they would give anything to have a baby, drink, or good nights sleep (not necessarily in that order). Whether such an extreme response is credible or not, sample means including these infinite responses are not defined. The potential for an infinite response or infinite value is a challenging aspect to all forms of preference-based measurement, and is acutely important in health valuation.

Currently, three tradeoff techniques dominate the literature, each of which involves varying quantities of life (i.e., time, risk, and persons). For example, the time trade-off technique (TTO) might ask whether the respondent prefers ten years in a disease state or eight years in optimal health. By raising and lowering the time in optimal health (a.k.a., quality-adjusted life years or QALYs), the interviewer can identify the respondent's indifference point (e.g., ten years in the health state equal to eight QALYs). A state may have an extremely negative value: a respondent may be indifferent between a minute with a disease and the loss of one QALY (one minute with disease equals negative one QALY), which implies that a year with disease is worth  $-525,949$  QALYs. Such an extreme TTO response would overwhelm typical summary measures. Therefore, this paper introduces an application of directional statistics in health valuation studies that may replace the more common practices (Dolan, 1997; Shaw, Johnson, & Coons, 2005).

The classical approach in health valuation remains highly controversial: (1) value is expressed as a ratio, representing the tradeoff between two goods (e.g.  $-1$  year /  $1$  minute =  $-525,949$  QALYs); (2) the summary measure is the mean ratio; and (3) because means with outliers behave badly, extreme ratios are arbitrarily transformed to make the estimates look more credible (Dolan, 1997; Lamers, 2007). We show that the assumption of angular error or "wavering preference" motivates the use of directional statistics as an alternative approach to ratio statistics, and negates the impetus toward arbitrary transformations.

## BACKGROUND

All trade-offs may be expressed using Cartesian coordinates  $(x,y)$ , where a person is indifferent between  $x$  and  $y$ . In time trade-off (TTO),  $x$  is time in a disease state and  $y$  is time in optimal health (a.k.a. quality-adjusted life years (QALYs)). Value may be expressed by the ratio,  $y/x$  (See Figure 1). For example, from a maximum choice of ten health years, a respondent may equate eight years in optimal health to ten years of disease time (Point A; value =  $8/10 = 0.8$ ). Or, a respondent may consider a scenario of two years of optimal health followed by eight years of disease, and equate it to "immediate death." Because the value of death is zero on a QALY scale, this response suggests the eight years of disease is equal to a loss of two years in optimal health (Point B; value =  $-2/8 = -0.25$ ). All TTO responses  $(x,y)$  may be arranged on the dashed line in Figure 1, and values on a QALY scale are bounded between one and negative infinity.

Using a sample of trade-off responses, the conventional approach to valuation is to estimate a ratio statistic,  $\mu$ , by minimizing the sum of squared error:

$$\underset{\mu}{\text{Min}} \frac{1}{N} \sum_{i=1}^N \varepsilon_i^2 \quad \text{where} \quad \varepsilon_i = y_i/x_i - \mu \quad (1)$$

Values,  $y_i/x_i$ , vary across individuals. Additive variation around a ratio statistic,  $\mu$ , may be expressed using an error term,  $\varepsilon_i$ , representing randomness in value and measurement error. Typically, additive error distributions have finite variance in addition to the expected value zero and independence. However, error in a ratio statistic can be poorly defined, because if one

or more  $x$  values are zero, the error becomes infinitely large. In the UK valuation of EQ-5D states, Dolan addressed the infinite variance problem by arbitrarily replacing all negative ratios (i.e., worse-than-death or WTD ratios) with  $y/10$  (Dolan, 1997). Because  $y$  in the TTO varies from negative 9.75 to 10 years, the range of the adjusted values is bounded, from - 0.975 to one.

Another concern with the application of ratio statistics is that  $x$  and  $y$  are not interchangeable. In other words,  $\mu(x,y)$  is not equal to the inverse of  $\mu(y,x)$ . This is particularly evident if one or more  $y$  values are zero. However, in a more general conjoint analysis, the tradeoff of  $x$  for  $y$  may or may not be equal to the inverse of the tradeoff of  $y$  for  $x$ , particularly in the case of complementary goods (e.g., shoe strings and shoes); however, this is an advantageous attribute in health valuation and other applications, like monetary exchanges (e.g., dollars for yen).

Drummond and colleague discuss similar difficulties in the estimation of incremental cost-effectiveness ratios (Drummond, 2005; Stinnett & Paltiel, 1997). On a cost-effectiveness plane, the  $y$ -axis reflects incremental costs ( $y$ ) and the  $x$ -axis reflects incremental effectiveness ( $x$ ). The convention is to divide the mean cost by the mean effectiveness (a ratio known as an incremental cost-effectiveness ratio or ICER) as an alternative to a ratio statistic.

The application of directional statistics in health valuation addresses the problems of extreme values and interchangeability, and motivates an estimator nearly identical to the ICER (i.e., ratio of means).

## METHODS

### Directional Statistics in Health Valuation

Every point in a Euclidean space can be uniquely mapped to a set of polar coordinates ( $\theta$ ,  $r$ ) described by an angle and a radius:

$$x=r*\cos(\theta) \quad y=r*\sin(\theta) \quad r=\sqrt{x^2+y^2} \quad \text{and} \quad \theta=\arctan(y/x)$$

Specifically, each ratio ( $y/x$ ) is the tangent of an angle,  $\theta$ . The radius,  $r$ , represents the size of the tradeoff. Instead of a ratio statistic as the value estimator, we propose estimating the tangent of the mean angle.

In Figure 1, we show that the QALY angles are bounded between 45 degrees and negative 90 degrees. For example, a non-trader's response of a negative infinite QALY value is a negative 90 degree QALY angle. Similarly, the value of optimal health (ratio = 1) is a 45 degree angle and the value of dead (0) is a zero degree angle.

Instead of expressing randomness in value,  $y_i/x_i = \mu + \varepsilon_i$ , we may express randomness in direction,  $\theta_i = \theta + \varepsilon_i$ . Direction error has been examined in many settings, such as adjustments on a dial, readings on a compass or clock, or the variability of wind directions or seasonality (Gao et al., 2006; Gao et al., 2005; Gao et al., 2002). As in the aforementioned examples, respondent preferences in health valuation may waver in a directional fashion (e.g., feeling up beat or dntrodden).

Our solution of changing the coordinate system so that problems can be circumvented (or calculations be made more easily) is commonly used in physics. For example, obtaining the equation of motion of a system of coupled oscillators can be done easier using Langrangian mechanics with spherical coordinates than using Newtonian mechanics with Cartesian coordinates (Landau & Lifshitz, 1989).

Because angles are bounded, directional statistics are finite by construction and interchangeable. However, two well known issues prevent the use of ordinary least squares (OLS) (equation 1) as a directional loss function for the estimation of mean angles: the crossover problem and circular variance. The crossover problem is related to the circular nature of angles. For example, on a compass, where north is zero degrees, the arithmetic mean of 45 degrees (northeast) and 315 degrees (northwest) is 180 degrees (south), not 0 degrees (north), even though zero may be a more accurate representation of central tendency. The potential of crossing over north prevents the use of arithmetic means in directional applications. The QALY angles lie between 45 degrees (the values of optimal health) and negative 90 degrees (value of negative infinity), not throughout the entire circle. Therefore, crossover (i.e., angles beyond 180 or negative 180 degrees) is not possible.

Because the sum of squared error does not represent circular variance, OLS (equation 1) is inappropriate to use as a directional loss function. The largest possible error in QALY angles is 145 degrees; yet, the OLS specification allows for error beyond 145 degrees, and the square of this error may reach beyond the crossover point (180 degrees). OLS is inappropriate for the estimation of a linear probability model for similar reasons.

In directional statistics, circular variance is represented by

$$1 - \sin(\widehat{\theta}) \sin(\theta) - \cos(\widehat{\theta}) \cos(\theta) \rightarrow \text{Min}_{\widehat{\theta}} \frac{1}{N} \sum_{i=1}^N 1 - \sin(\widehat{\theta}) \sin(\theta_i) - \cos(\widehat{\theta}) \cos(\theta_i) \tag{2}$$

Mean angle,  $\widehat{\theta}$ , is the estimate that minimizes circular variance, which is a directional loss function analogous to OLS (equation 1). Mardia and Jupp refer to this measure of dispersion as one minus the mean resultant length,  $\bar{R}$  (Mardia & Jupp, 2000). Unlike the error in ratio statistics, each element in the circular variance expression is finite, ranging from zero to two, with an overall mean ranging between zero and one. If the angles are widely dispersed (i.e., discordance in health state value), circular variance approaches one,  $\bar{R} \cong 0$ , and if the angles are concentrated (i.e., concordance in health state value), circular variance approaches zero,  $\bar{R} \cong 1$ .

To clarify the estimator of the tangent, we take the derivative of equation 2 and set it to zero:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N -\cos(\widehat{\theta}) \sin(\theta_j) + \sin(\widehat{\theta}) \cos(\theta_j) &= 0 \\ \cos(\widehat{\theta}) \left( \frac{1}{N} \sum_{i=1}^N -\sin(\theta_j) + \tan(\widehat{\theta}) \frac{1}{N} \sum_{i=1}^N \cos(\theta_j) \right) &= 0 \\ \tan(\widehat{\theta}) &= \frac{\frac{1}{N} \sum_{i=1}^N \sin(\theta_j)}{\frac{1}{N} \sum_{i=1}^N \cos(\theta_j)} \\ \tan(\widehat{\theta}) &= \frac{\frac{1}{N} \sum_{i=1}^N \frac{y_i}{r_i}}{\frac{1}{N} \sum_{j=1}^N \frac{x_j}{r_j}} \end{aligned}$$

The tangent of the mean angle is the mean of y/r over the mean of x/r, where r is the radius. In Figure 1, each TTO response has a radius as measured from the distance to the origin. If all responses (x,y) were rescaled by dividing by their radii, they would lie along the semi-circular line. The tangent of the mean angle would be the mean of the rescaled y over the mean of the rescaled x. In other words, the mean angle estimator ignores the distance from the origin of each response.

Instead, radii may be included in the loss function as weights for each element of equation two. It follows that the tangent of a radially weighted mean angle is the mean of  $y$  over the mean of  $x$ :

$$\widehat{\min}_{\theta_r} \frac{1}{N} \sum_{i=1}^N r_i \left( 1 - \sin(\theta_i) \sin(\widehat{\theta}_r) - \cos(\theta_i) \cos(\widehat{\theta}_r) \right) \rightarrow \tan(\widehat{\theta}_r) = \frac{1}{N} \sum_{i=1}^N y_i / \frac{1}{N} \sum_{j=1}^N x_j \quad (3)$$

Radial weighting the loss function suggests that angular error far from the origin is more important than error near the origin. In valuation, tradeoffs with lengthy radii may be given more weight, because greater quantities are involved. For example, a monetary exchange involving millions of Euros may receive greater attention than a typical money exchange at an automatic teller machine. On the other hand, in trade-off response, which represents a single respondent's valuation of a single state (i.e., one person, one vote), variability in the radii (see Figure 1) is viewed as an artifact of the experimental design, and the mean angle removes this arbitrary noise. The point (9,10) is farther from the origin than the point (8,10), but this does not necessarily suggest that it is more or less important.

In health valuation, directional statistics are appealing for their technical simplicity and plain intuition (i.e., individual preferences waver). Instead of the ratio statistic (i.e., the mean of  $y/x$ ), the approach entails a ratio of means,  $\bar{y} / \bar{x}$ . Radially weighted or not, the mean of  $x$  is non-zero by construction; therefore, the directional statistics may be more robust than their ratio counterparts. If  $x$  and  $y$  were switched, the resulting estimate would be the inverse of the original (i.e., interchangeability). When Dolan replaced WTD responses with  $y/10$ , the adjusted ratio statistic became  $\bar{y} / 10$ , which is similar to the radially weighted estimator,  $\bar{y} / \bar{x}$  (Dolan, 1997). Although Dolan's transformation has no theoretical basis, estimates under the classical approach approximate those based on directional statistics by construction.

### Circular Regression

Valuation studies typically examine tradeoffs between hypothesized health scenarios to predict the values of scenarios that were not directly incorporated into the sample. Out-of-sample predictions can be accomplished using a linear combination of state-specific variables,  $Z'\beta$ , known as a multi-attribute utility (MAU) regression model. Using OLS (equation 1), the classical approach is to estimate the MAU regression model,  $y_i/x_i = Z_i'\beta + \varepsilon_i$ , where the dependent variable is the ratio,  $y/x$ . To improve the face validity of these predictions, Dolan arbitrarily replaced the dependent variable with  $y/10$ .

The circular regression approach is to estimate a linear MAU model by minimizing circular variance (equation 2), where  $\theta_i = \arctan(y_i/x_i)$  and  $\widehat{\theta} = \arctan(Z_i'\widehat{\beta})$ . Similarly, the radially weighted directional loss function (equation 3) may be minimized to estimate  $\widehat{\theta}_r = \arctan(Z_i'\widehat{\beta}_r)$ . The MAU regression coefficients,  $\beta$  and  $\beta_r$ , are on the same scale as the ratio statistic estimates; yet, the circular regression approach avoids the problems of ratio statistics and the arbitrary transformations of WTD responses.

### The Measurement and Valuation of Health Study in the United Kingdom

To demonstrate the application of directional statistics in health valuation, we examine data from the seminal Measurement and Valuation of Health Study (Dolan, 1997; Gudex, 1994). In 1993, the University of York administered 3395 interviews with a response rate of 64% and collected values of 42 EQ-5D health states and the state of unconsciousness. During the TTO exercise, respondents placed a value on up to 13 states. As mentioned earlier, the MVH protocol

bounded the lower end of the loss in years to be greater than negative 9.75 (See Figure 1); therefore, the ratio,  $y/x$ , is bounded between 1 and -39 (or  $-9.75/0.25$ ).

For the TTO analytical sample ( $N=3,355$ ), respondents were excluded (1) if only one or two states were valued (other than 11111, “immediate death,” and “unconscious”); (2) if all states were given the same value; or (3) if all states were valued worse than “immediate death.” The three criteria motivated the exclusion of 1.2% of the TTO respondents. Across the 3,355 respondents, each of the 39,673 TTO responses described an equivalence of time in optimal and non-optimal health, BTD (10,y) or WTD (10+y, y), where y is time in optimal health between 10 and negative 9.75 years.

In this analysis, the values of the 42 hypothesized EQ-5D states were estimated using ratio statistics with and without Dolan's transformation of WTD responses, and using directional statistics with and without radial weights. This allowed the comparison of the four methods (i.e., mean ratio, Dolan adjusted, Unweighted, and Radially-Weighted) without the distraction of state-specific attributes. Likewise, four MAU regression models were estimated to predict the values of the 243 EQ-5D states.

For both the 42 state values and the regression coefficient, 95% confidence intervals were estimated using the percentile method by applying bootstrap sampling with respondent-specific cluster replacement. For each iteration of the bootstrap, a sample of respondents was extracted from the analytical sample with replacement and the analysis was re-run with the bootstrap sample. After 1,000 iterations, the parameter estimates were sorted and the top and bottom 24 estimates of each parameter removed. The 25<sup>th</sup> and 975<sup>th</sup> estimates represented the 95% confidence interval under the percentile bootstrap approach (Efron & Tibshirani, 1993).

In complement to visual inspection, concordance between the predictions made in this study was measured using Lin's coefficient of agreement and mean absolute difference. Because of its prominence in the literature, Dolan's published value set was compared to these regression predictions. Using the same source data and variables as the original analysis of the MVH data, ratio statistic estimates with Dolan's transformation of WTD responses were nearly identical in this analysis to published estimates. Minor deviations between the published and the re-estimated values may be attributable to differences in sample selection criteria.

Like the original analysis of the MVH data, the MUA regression model includes twelve indicator variables: five for second level domains, five for the third level domains, one for any second or third level domains (i.e., constant); and one for any third level domains (i.e. N3). The EQ-5D descriptive system has five dimensions (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) each with three possible levels (Szende et al., 2007). The MUA regression model captures the detrimental effects of each level on each domain as well as the multiplicative effects of any one second or third level domains (i.e., constant), and of any one third level domains in the state vector (i.e., N3). All database work was conducted on SAS 9.1, and the analyses were conducted using Stata 10 (SAS, 2007; StataCorp, 2007).

## ANALYSIS

After stratifying the analytic sample into 42 state-specific subsamples, we estimated the mean ratio with and without Dolan's transformation of WTD responses (i.e.,  $\frac{1}{N} \sum y_i/x_i$  and  $\bar{y} / 10$ ) and the tangent of the mean angle with and without radial weights (i.e.,  $\bar{y}_r/\bar{x}_r$  and  $\bar{y} / \bar{x}$ ). Without the Dolan transformation, the mean ratio is significantly positive for only ten out of the 42 states, which suggests that few states are better than “immediate death”. These estimates clearly lack face validity, and motivate the Dolan transformation. Transformed estimates are significantly greater than or equal to mean ratios for all 42 states, arbitrarily increasing values.

Of the 42 states, 28 of the transformed estimates are significantly positive and 14 significantly negative, suggesting a third of the states are worse than “immediate death”. Because the Dolan transformation is the conventional approach to health valuation, these estimates (i.e., Dolan ratios) are compared with directional results.

Based on Figure 2, directional statistics produce values similar to the Dolan ratios. As a measure of concordance with the arbitrarily adjusted estimates, Lin's coefficient of agreement is 0.916 for the tangent of the mean angles and 0.954 for the radially weighted estimates. Similarly, the absolute mean difference is 0.136 for the unweighted and 0.067 for the weighted.

While the estimates are similar, the unweighted estimates are significantly less than the Dolan ratios for all 42 states (Figure 2). The weighted estimates are more balanced; significantly less than the Dolan ratios for 26 states and greater than for 14 states. If the purpose is to produce estimates similar to the Dolan ratio predictions, the tangent of the radially weighed angle is a preferred estimator.

Figure 2 further illustrates the negative relationship between the Dolan ratios and the angle-based QALY values. For BTD states, the differences between the Dolan ratios and the directional estimates appear small. For WTD states, the difference increases as states grow more severe. For example, the Dolan ratio of the “pits” state (33333) is X, and the weighted estimate is Y.

### Health Valuation of the Entire EQ-5D Descriptive System

To assign values to all 243 possible EQ-5D health states, we estimated four regression models (Table 1). Each coefficient reflects a decrement from optimal health (1.00); therefore, based on the 95% bootstrap confidence intervals, it is expected to be significantly negative. For example, the Dolan ratio coefficient of any one second or third level domain is -0.086, which suggests that non-optimal health states have a maximum value of 0.914 (or 1-0.086). This decrement is known as the non-optimal gap.

The first five coefficients represent the decrement associated with “some problems” on each of the five domains. For these coefficients, the 95% confidence intervals overlap. The second set of coefficients represents decrements associated with “severe problems.” With the exception of the N3 coefficient, the Dolan ratio coefficients are significantly lower than the directional coefficients. The value of the N3 coefficient, being larger in the Dolan ratio model (-0.279) than in the directional models (-0.125 and -0.028), suggests that the directional models better differentiate the third level domains than the Dolan ratio model.

The replication of the Dolan ratio model is nearly identical to the published estimates (Figure 3). Lin's coefficient is 0.999, and the mean absolute difference across the 243 predicted values is 0.006. The small difference is likely attributable to rounding error and changes in the sample selection criteria. Figure 3 illustrates the relationship between the published Dolan estimates and the angle-based predictions, including predictions for all 243 EQ-5D states (Dolan, 1997). A negative relationship is illustrated where the greatest difference appears in the prediction of WTD values. For the 243 predictions, Lin's coefficient of agreement between Dolan's values and the unweighted values is 0.85, and the mean absolute difference is 0.164. Greater agreement is found in the weighted estimates, where Lin's coefficient is 0.922 and the mean absolute difference is 0.109. For reference, Figure 4 is a histogram of angular error for the circular regression models with or without radial weights.

## DISCUSSION

In this paper, we introduce the concept of wavering preferences, and two directional statistics for use in the valuation of health states (i.e.,  $\bar{y}_r/\bar{x}_r$  and  $\bar{y}/\bar{x}$ ). Each estimator addresses well known issues in the ratio statistics, specifically infinite values and interchangeability, and negates the impetus behind the transformation of outlying responses (e.g., Dolan or Shaw's transformation of WTD responses). The estimator is nearly identical to an incremental cost-effectiveness ratio (ICER). The resulting predictions are similar to those commonly applied in health policy; however, differences occur in the more severe health states. We focus on health valuation; however, multiple areas of conjoint analysis in health and medicine may benefit from this approach, particularly those where tradeoffs seem unfathomable.

Given that the QALY scale is bounded between one and negative infinity, QALY angles are bounded between 45 degrees and negative 90 degrees. The directional estimator does not impose these bounds, which leads to two possible limitations. First, the predicted angles may be outside the QALY scale, which is similar to the problems faced in linear probability models. The ratio of means,  $\bar{y}/\bar{x}$ , are naturally bounded to the interval, but out-of-sample predictions may be off the scale. This is unlikely to occur at the upper bound where 45 degrees is optimal health, because all descriptive systems describe decrements from this point. WTD values may seem more likely to extend past negative 90 degrees, but few states are WTD. A second limitation is that the confidence intervals may span outside the QALY scale. To address this limitation, we apply bootstrap techniques to estimate the confidence intervals around estimates instead of assuming symmetric standard error. Because bootstrap intervals are empirical and rely on resampled predictions, the confidence intervals remain within the QALY scale.

Alternative statistical methods in health valuation have been proposed to analytically accommodate infinite ratios, all of which have underlying assumptions with arbitrary elements. The most common is the transformation of WTD responses. Lamers and colleagues investigated three such transformations: 1) the monotonic transformation,  $y/10$ , proposed by Patrick and used by Dolan; 2) the linear transformation,  $(y/x)/39$ , proposed by Shaw and colleagues; and 3) truncation at -1 (Dolan, 1997; Lamers, 2007; Patrick et al., 1994; Shaw et al., 2005). Lamers shows that each renders a different value set, and all transformations lack a sound theoretical underpinning.

A second class of alternative methods involves changing the estimator, not the data. In the current paper, we recommend the use of direction statistics; however, Craig and Busschbach recommend regressing  $y$  on  $x$ , using a coefficient, instead of a ratio statistic, as the estimator (Craig & Busschbach, 2009). More recent work has investigated changing the measure of central tendency: instead of mean ratio, median or mode ratios may be estimated (Shaw et al., 2007). Median and mode statistics mitigate the effects of potentially infinite distribution tails, but are less relevant for economic evaluations.

Choosing directional statistics over Craig and Busschbach's coefficient approach may appear arbitrary; nevertheless, each has a clear utility framework (i.e., wavering and episodic utility) (Craig & Busschbach, 2009). No theoretical framework has yet been proposed to motivate the manipulation of data or the use of medians or modes for decision analyses, so these more pragmatic alternatives may be less justified.

Changing the estimator does not resolve issues inherent to tradeoff experimental protocols. TTO responses are collected on two scales, one for BTD responses and another for WTD responses (Gudex, 1994). Scale separation may psychometrically influence TTO responses, which is not addressed by the proposed directional statistics. Secondly, we examine the value of a health state by varying time as a quantity of life, not risk or persons. Even though the



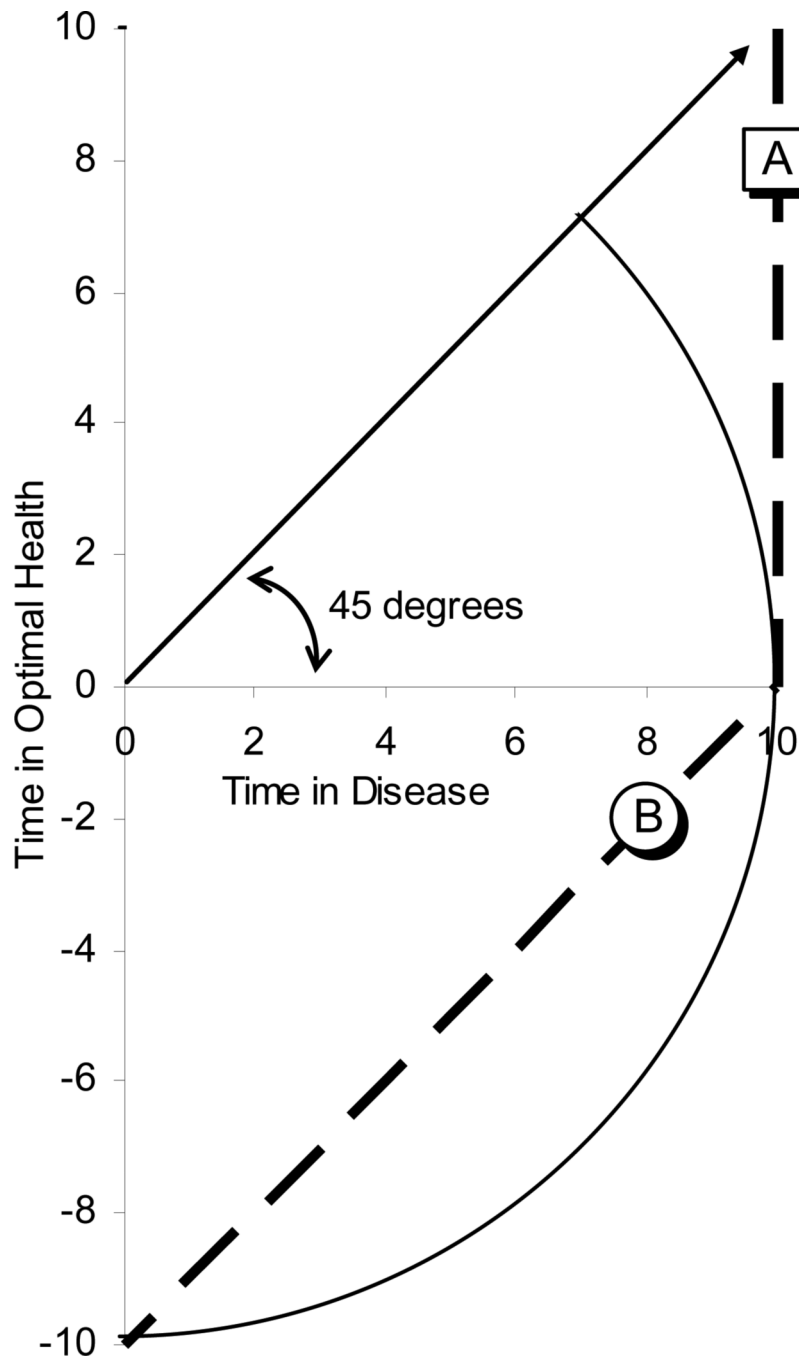
problem that we present in this paper is essentially two dimensional, this does not mean that the use of directional statistics is limited to two dimensional problems. In principle the methodology can be extended to include three or more dimensional problems, just like in physics (e.g. in relativistic mechanics a four dimensional coordinate system is typically used) (Bleichrodt, 2002; Craig, 2009; Craig et al., 2009). Lastly, the TTO task involves only gains in time. Prospect theory suggests that respondents value losses distinctly from gains, and adjusting for these differences would be analogous to adding a reverse gear to the directional approach (Oliver, 2003; van Osch & Stiggelbout, 2008).

The tangent of the radially weighted mean angle,  $\bar{y} / \bar{x}$ , provides consistent estimates without the arbitrary transformation of WTD responses, and the estimator has a clear underlying theoretical framework (i.e., waning preferences). Its predictions are nearly identical to Dolan's estimates, except that they have a wider range. To understand this difference, it is noteworthy that the two estimators,  $\bar{y} / 10$  and  $\bar{y} / \bar{x}$ , are the same except for the difference between  $\bar{x}$  and ten. Because time in disease ( $x$ ) is ten years or less, no simulation is required to show that the proportional difference between the estimates is always  $\bar{x} / 10$  by construction. The more difficult questions concern the implications of this wider range in QALY estimates for economic evaluations.

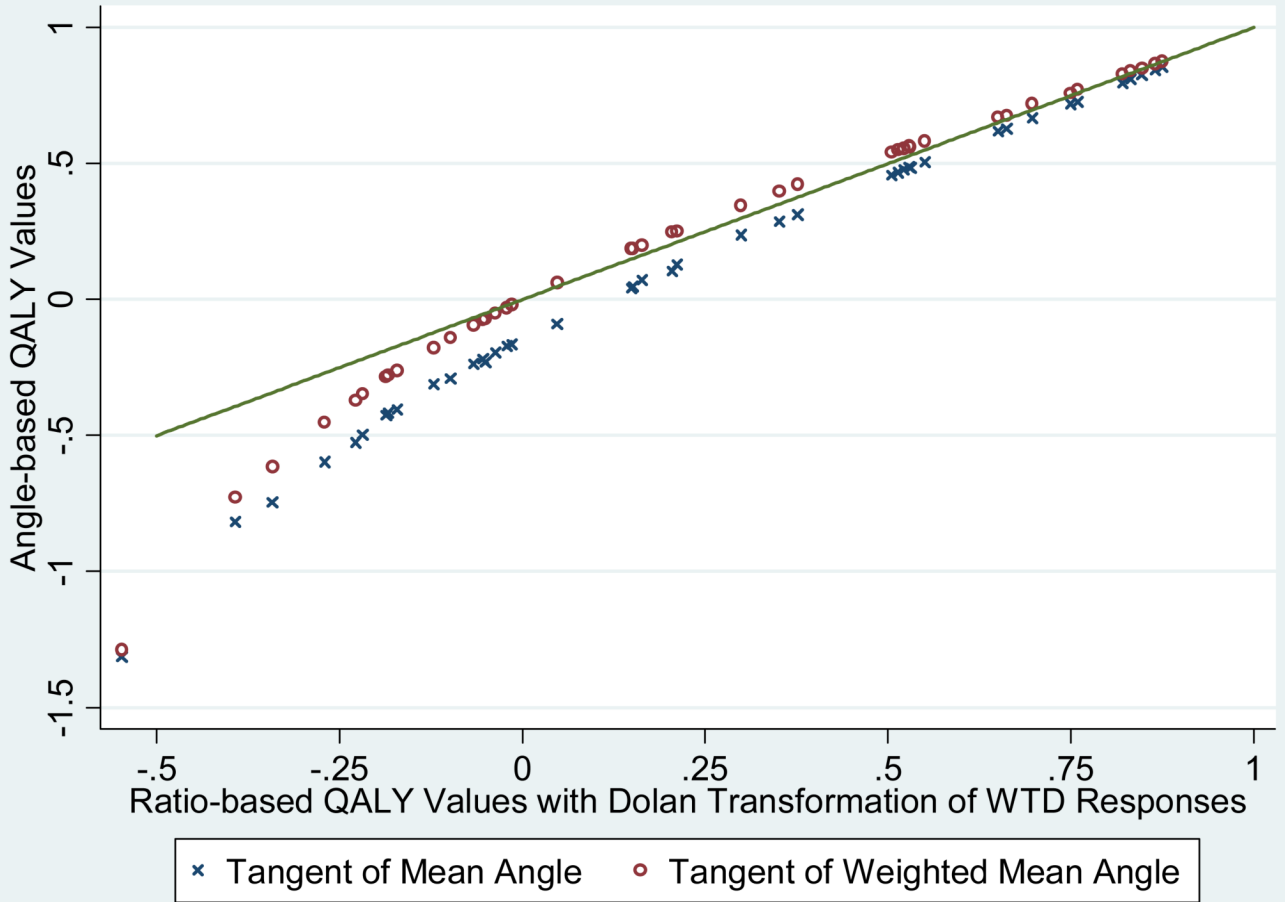
## REFERENCES

- Bleichrodt H. A new explanation for the difference between time trade-off utilities and standard gamble utilities. *Health Econ* 2002;11(5):447–456. [PubMed: 12112493]
- Craig B. The duration effect: a link between TTO and VAS values. *Health Econ* 2009;18(2):217–225. [PubMed: 18351621]
- Craig B, Busschbach J. The episodic random utility model unifies time trade-off and discrete choice approaches in health state valuation. *Popul Health Metr* 2009;7(1):3. [PubMed: 19144115]
- Craig B, Busschbach J, Salomon J. Ranking, Time Trade-Off and Visual Analogue Scale Values for EQ-5D Health States. *Medical Care*. 2009 forthcoming.
- Dolan P. Modeling valuations for EuroQol health states. *Medical Care* 1997;35(11):1095–1108. [PubMed: 9366889]
- Drummond, M. *Methods for the economic evaluation of health care programmes*. Oxford University Press; Oxford, New York: 2005.
- Efron, R.; Tibshirani, B. *An Introduction to the Bootstrap*: Chapman and Hall. 1993.
- Gao F, Chia K, Krantz I, Nordin P, Machin D. On the application of the von Mises distribution and angular regression methods to investigate the seasonality of disease onset. *Stat Med* 2006;25(9):1593–1618. [PubMed: 16382488]
- Gao F, Nordin P, Krantz I, Chia K, Machin D. Variation in the seasonal diagnosis of acute lymphoblastic leukemia: evidence from Singapore, the United States, and Sweden. *Am J Epidemiol* 2005;162(8):753–763. [PubMed: 16135507]
- Gao F, Seah S, Foster P, Chia K, Machin D. Angular regression and the detection of the seasonal onset of disease. *J Cancer Epidemiol Prev* 2002;7(1):29–35. [PubMed: 12369603]
- Gudex, C. *Report of the Centre for Health Economics*. University of York; York, United Kingdom: 1994. *Time Trade-Off User Manual: Props and Self-Completion Methods*..
- Lamers L. The transformation of utilities for health states worse than death: consequences for the estimation of EQ-5D value sets. *Med Care* 2007;45(3):238–244. [PubMed: 17304081]
- Landau, L.; Lifshitz, E. *Mechanics*. Pergamon Press; Oxford, New York: 1989.
- Mardia, K.; Jupp, P. *Directional statistics*. J. Wiley; Chichester, New York: 2000.
- Oliver A. The internal consistency of the standard gamble: tests after adjusting for prospect theory. *J Health Econ* 2003;22(4):659–674. [PubMed: 12842320]
- Patrick D, Starks H, Cain K, Uhlmann R, Pearlman R. Measuring preferences for health states worse than death. *Med Decis Making* 1994;14(1):9–18. [PubMed: 8152361]

- SAS. SAS 9.1 for Windows. SAS Institute Inc.; Cary, NC, USA: 2007.
- Shaw J, Johnson J, Coons S. US valuation of the EQ-5D health states: development and testing of the D1 valuation model. *Med Care* 2005;43(3):203–220. [PubMed: 15725977]
- Shaw, J.; Shengsheng, Y.; Chen, S.; Iannacchione, V.; Johnson, J.; Coons, S. There's nothing “Average” about these weights: Development and testing of a median model of US EQ-5D Health State Preferences.. 14th Annual Meeting of the International Society for Quality of Life Research.; Toronto, ON. 2007.
- StataCorp. Stata Statistical Software: Release 10. StataCorp LP; College Station, Texas, USA: 2007.
- Stinnett A, Paltiel A. Estimating CE ratios under second-order uncertainty: the mean ratio versus the ratio of means. *Med Decis Making* 1997;17(4):483–489. [PubMed: 9343807]
- Szende, A.; Oppe, M.; Devlin, N., editors. EQ-5D value sets : inventory, comparative review, and user guide. Springer; Dordrecht: 2007.
- van Osch S, Stiggelbout A. The construction of standard gamble utilities. *Health Econ* 2008;17(1):31–40. [PubMed: 17410521]



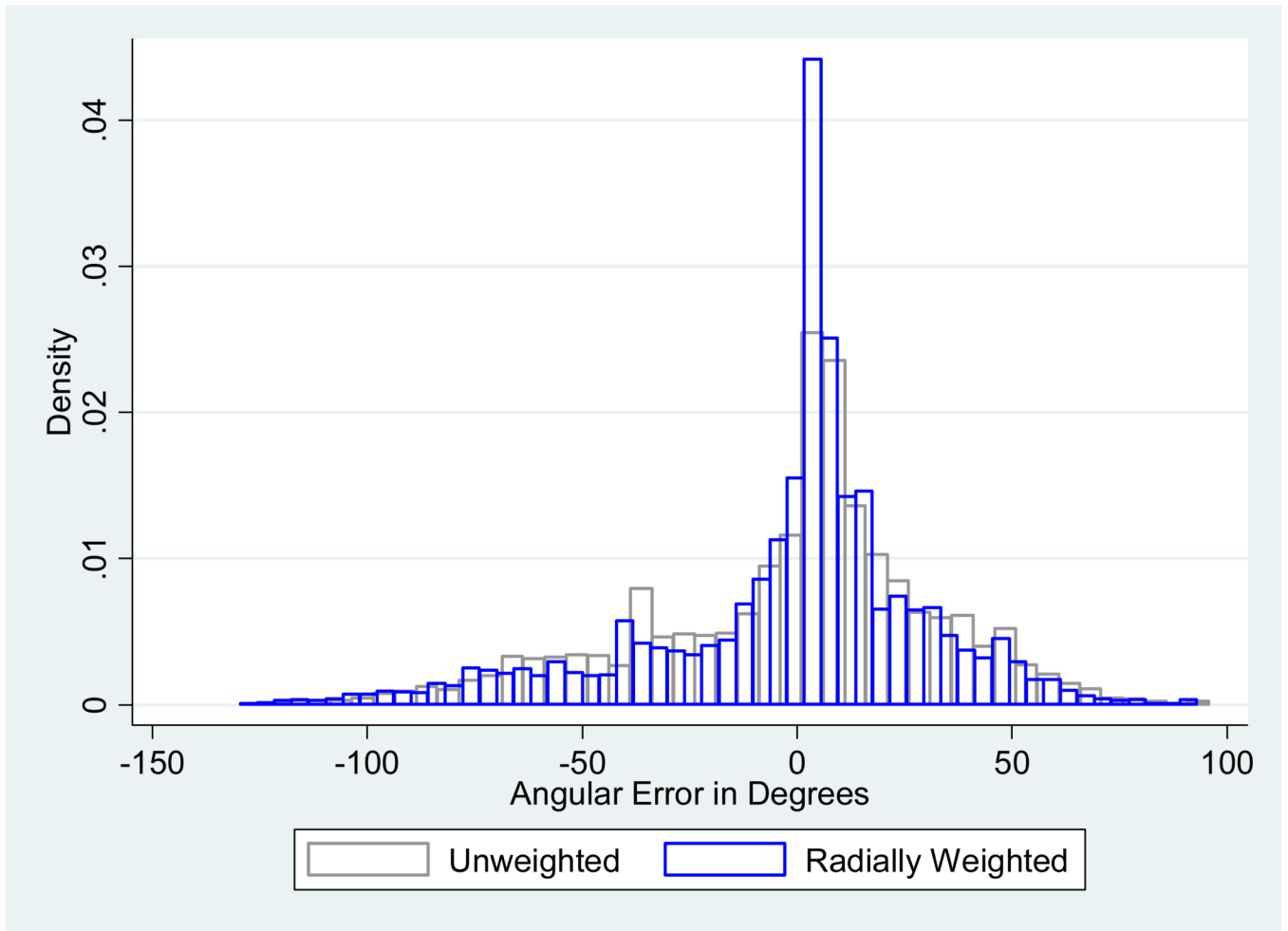
**Figure 1.**  
Time Trade-off Responses



**Figure 2.** Comparison of Angle-based and Ratio-based QALY values for the 42 Hypothesized EQ-5D Health States



**Figure 3.** Comparison of Angle-based and Ratio-based QALY values for All 243 possible EQ-5D Health States



**Figure 4.**  
Histogram of Angular Error for Circular Regression Models

**Table 1**

Multi-Attribute Utility Regression Models for EQ-5D Health States

N= 3,355 respondents with 39,673 responses

State Attributes	Mean Ratio *				Tangent of Mean Angle				
	Untransformed		Transformed		Unweighted		Radially Weighted		
	Coef.	95% C.I.	Coef.	95% C.I.	Coef.	95% C.I.	Coef.	95% C.I.	
Mobility, 2	-0.234	-0.378	-0.069	-0.081	-0.055	-0.073	-0.038	-0.040	-0.025
Self-Care, 2	-0.085	-0.227	0.051	-0.116	-0.094	-0.115	-0.081	-0.082	-0.068
Usual Activity, 2	-0.246	-0.372	-0.114	-0.034	-0.022	-0.065	-0.031	-0.050	-0.037
Pain/Discomfort, 2	-0.240	-0.400	-0.095	-0.131	-0.108	-0.142	-0.105	-0.106	-0.090
Anxiety/Depression, 2	-0.213	-0.346	-0.081	-0.071	-0.082	-0.104	-0.072	-0.087	-0.073
Mobility, 3	-2.627	-2.900	-2.365	-0.311	-0.328	-0.293	-0.499	-0.498	-0.469
Self-Care, 3	-1.614	-1.837	-1.405	-0.217	-0.232	-0.201	-0.327	-0.323	-0.301
Usual Activity, 3	-1.202	-1.450	-0.977	-0.084	-0.102	-0.068	-0.183	-0.204	-0.181
Pain/Discomfort, 3	-2.710	-2.962	-2.468	-0.374	-0.390	-0.358	-0.590	-0.575	-0.549
Anxiety/Depression, 3	-2.050	-2.284	-1.800	-0.234	-0.250	-0.219	-0.386	-0.359	-0.370
Any 2's or 3's	-0.073	-0.163	0.029	-0.086	-0.075	-0.114	-0.129	-0.090	-0.078
Any 3's	0.834	0.583	1.075	-0.279	-0.297	-0.260	-0.125	-0.028	-0.005

\* In the TTO, better than death (BTD) response is (10,y) and worse than death (WTD) response is (10+y, y) where y is years in optimal health equal to x years in the health state. Dolan transformed the WTD responses from (10+y,y) to (10,y), arbitrarily inflating the ratios, y/x.