Economics and Human Biology 15 (2014) 41–46

Contents lists available at ScienceDirect

Economics and Human Biology

journal homepage: http://www.elsevier.com/locate/ehb

Are people consistent when trading time for health?

Adam Oliver [a,*], Jonathan Wolff [b,1]

[a] Department of Social Policy, London School of Economics, Houghton Street, London WC2A 2AE, UK
[b] Department of Philosophy, University College London, Gower Street, London WC1E 6BT, UK

ARTICLE INFO

Article history:
Received 25 September 2013
Received in revised form 29 April 2014
Accepted 1 May 2014
Available online 28 May 2014

Keywords:
Time trade-off
Procedural invariance
Loss aversion

ABSTRACT

The conventional, or standard, time trade-off (TTO) procedure, which is used to elicit the values that people place on health states that are in turn required to calculate quality adjusted life-years (QALYs), asks respondents to trade off fewer life years for better health. It is possible to reverse the procedure to ask respondents to trade off less health for more life years. Theoretically, these two procedures should generate the same TTO values for any given health state. This article reports that for health states defined by differing frequencies of migraine attack, the standard TTO gives health state values that are significantly higher than those given by the reverse TTO. The observed systematic procedural invariance, which substantiates some previous findings reported in the literature and is consistent with a loss aversion effect, challenges the validity of the TTO for generating reliable valuations of health states.

1. Introduction

Cost-utility analysis (CUA), a special form of cost-effectiveness analysis (CEA) that typically defines outcomes in terms of quality adjusted life-years (QALYs), has been explicitly and formally used to inform health care resource allocation in England and Wales (and elsewhere) for many years, and since 1999 has been recommended by The National Institute for Health and Care Excellence (NICE). While increasingly influential internationally, CUA has not escaped criticism on a number of levels.

One type of critic questions the justification of the CUA approach, ruling it morally unacceptable to deny someone a beneficial treatment on cost grounds (Harris, 1987). More moderate critics argue that this form of health economic evaluation is not without merit but alone is not enough, and an appeal to a set of further values or factors over and above straightforward QALY maximisation is also essential (Orr et al., 2011). For example, some argue that a severe illness should have priority over a less severe one, even if treating them is equally cost effective. Such criticisms have been broadly accepted by NICE in the development of its methodology (NICE, 2008). A third type of critic asks whether we do, in fact, have a reliable and robust method for valuing health states. If, for example, different approaches yield strikingly different valuations then an intervention that is ruled cost-effective with use of one method could fail to be so with use of another, a point to which we will later return. If it turns out that our best methods are flawed, yielding inconsistent results when they are examined carefully, then the economic analysis may rest on shaky foundations.

In this article we are concerned with objections of the third type. We do not take a stand on the morality of CUA, QALYs or the need to incorporate further values. Rather, we look at the most commonly recommended and used method for eliciting health state values, the time trade-off (TTO), and consider a question that has, as detailed in the next section, also been addressed, albeit with mixed findings, by others: i.e. does the TTO method yield inconsistent results

* Corresponding author. Tel.: +44 020 7955 6471.
E-mail addresses: a.j.oliver@lse.ac.uk (A. Oliver), j.wolff@ucl.ac.uk (J. Wolff).
[1] Tel.: +44 020 7679 3067.

http://dx.doi.org/10.1016/j.ehb.2014.05.001
1570-677X/

depending on how the questions are asked? If evidence of systematic inconsistency (i.e. unidirectional inconsistency, such that one method shows a strong tendency to generate higher values than another method) is observed, then the challenge to the validity of the TTO cannot legitimately be attributed to random error.

2. Literature and hypotheses

For the computation of QALYs, numerical values for health states are needed. As aforementioned, in the area of health economic evaluation, the TTO is a commonly used and broadly accepted method by which to elicit such values. Borrowing notation from Bleichrodt (2002), the TTO is used to elicit the duration of T_2 that yields indifference between (Q_1, T_1) and (FH, T_2), where Q_1 is a health state typically intermediate in value between full health and death, T_1 is a pre-given number of years, and FH is full health. Utility is assumed to be linear in duration; i.e. $U(Q_1, T)$ is assumed to be equal to $H(Q_1)T$, where U is a utility function over outcomes and H is a utility function over health status. Given this assumption, indifference in the TTO question gives $H(Q_1)T_1 = H(\text{FH})T_2 = T_2$ (since $H(\text{FH}) = 1$); hence $H(Q_1) = T_2/T_1$, which is the formula typically used for eliciting TTO health state values.

Linear utility for duration is a strong assumption that is sometimes violated (Miyamoto and Eraker, 1985). If utility for duration is concave, for which positive time preference is a special case, the TTO values will be biased downwards (Johannesson et al., 1994; Dolan and Jones-Lee, 1997). Conversely, TTO values can also be biased upwards by scale compatibility (Bleichrodt, 2002), which, in the context of the TTO, is what Spencer (2003) categorises as a strong preference for longevity. According to Tversky et al. (1988), scale compatibility occurs when the response mode focuses attention on the compatible attributes of an option (e.g. monetary valuation tends to focus respondents' attention on money outcomes). In the TTO, duration is used as the response scale; therefore, a respondent who exhibits scale compatibility will place a strong emphasis on years of life and will therefore focus on securing longevity rather than better health. The respondent's indifference value for T_2 will thus be high, causing an upward bias on the TTO values.

Loss aversion can also upwardly bias TTO values when the TTO is presented in its standard format. Loss aversion is the observation that the disutility that people feel towards losses against a reference point is substantially greater than the utility that they feel towards gains of the same magnitude; indeed, for small to moderate money losses and gains, the ratio of disutility to utility is approximately 2–1 (Tversky and Kahneman, 1991), a finding that cannot be explained by standard economic theory. Loss aversion is an explanation for the endowment effect, or, in other words, the observation that people's maximum willingness to pay for a particular good tends to be far less than the minimum money amount that they are willing to accept to sell that same good (Kahneman et al., 1990). In the standard TTO the respondent is first presented with (Q_1, T_1), and this may consequently form the reference point. If so, the respondent is essentially being asked to trade off the gain in health status from Q_1 to full health against the loss in duration from T_1 to T_2. Again following Bleichrodt (2002), let T_2 be the respondent's answer in the absence of loss aversion, such that the utility of the gain in health status exactly offsets the disutility of the T_1 to T_2 loss in life years. Now assume that the respondent is loss averse, so that the loss in life years, $(T_1 - T_2)$, gets more weight than the gain in health status, $(\text{FH} - Q_1)$. In these circumstances, the respondent will prefer (Q_1, T_1) over (FH, T_2). For indifference, T_2 would need to increase, to T'_2, for instance. Thus, with loss aversion, the TTO value T'_2/T_1 exceeds the TTO value in the absence of loss aversion (T_2/T_1): hence, loss aversion upwardly biases the standard TTO values.

If the TTO question is instead asked in a reversed order – i.e. if respondents are asked how many years T_1 in Q_1 would make them indifferent to a pre-given T_2 years in full health – then the reference point may shift to (FH, T_2). Consequently, the respondent would trade off the gain in duration, T_2 to T_1, against the loss of health status, FH to Q_1, and loss aversion would predict a downward bias in the TTO values. This is because T_1 would have to be higher than in the absence of loss aversion in order to balance the loss in health status (Bleichrodt, 2002; Spencer, 2003). Thus, under loss aversion, the standard TTO would generate a higher value for any particular health state than the reverse TTO, a violation of procedural invariance that would challenge the validity of the instrument.

Spencer (2003) also speculated that if health state values are affected by feelings of a maximum endurable time (MET) – that is, if a health state is so bad, then beyond a particular point in time a person would prefer to die rather than continue to endure the health state – then the reverse TTO will generate values that are higher than those given by the standard TTO. This is because in the reverse TTO the respondent will not extend by much the time lived in the health state that is being valued, which will lead to a high TTO value. However, in the standard TTO, the respondent will accept a small amount of time in full health for indifference with the time spent in the intermediate health state, which will push the value downwards. MET preferences should not be influential if the evaluated health state is tolerable.

Scale compatibility should also lack influence when testing the difference between standard and reverse TTO values because the same response mode – i.e. years of life – is used in both frames, and thus, if the effect is at work, the respondents' attention will always be drawn towards life years rather than the health state. However, discounting may have an effect that is albeit ambiguous because the direction (positive, negative or neutral) of discounting is unknown, and may differ across respondent samples and/or according to the questions posed. It is noteworthy that even after attempting to control for discounting, Attema and Brouwer (2008) observed substantial differences in standard versus reverse TTO values, in the direction predicted by loss aversion, although the same authors in a different study reported in 2012 found that, over five tests, the standard TTO generated significantly different values to the reverse TTO only once, with and without controlling for discounting.

Spencer (2003) also tested whether values elicited with the standard TTO exceed those elicited with the reverse TTO,