HEALTH ECONOMICS
OUTCOME MEASUREMENT
Health Econ. **11**: 447–456 (2002)
Published online 4 March 2002 in Wiley InterScience (www.interscience.wiley.com). **DOI:**10.1002/hec.688

A new explanation for the difference between time trade-off utilities and standard gamble utilities

Han Bleichrodt*
iMTA/iBMG, Erasmus University, Rotterdam, The Netherlands

Summary

This paper gives a new explanation for the systematic disparity between standard gamble (SG) utilities and time trade-off (TTO) utilities. The common explanation, which is based on expected utility, is that the disparity is caused by curvature of the utility function for duration. This explanation is, however, incomplete. People violate expected utility and these violations lead to biases in SG and TTO utilities. The paper analyzes the impact on SG and TTO utilities of three main reasons why people violate expected utility: probability weighting, loss aversion, and scale compatibility. In the SG, the combined effect of utility curvature, probability weighting, loss aversion, and scale compatibility is an upward bias. In the TTO these factors lead both to upward and to downward biases. This analysis can also explain the tentative empirical finding that the TTO better describes people's preferences for health than the SG. Copyright © 2002 John Wiley & Sons, Ltd.

JEL classification: I10; D81

Keywords health utility measurement; standard gamble; time trade-off; probability weighting; loss aversion; scale compatibility

Introduction

The standard gamble (SG) and the time trade-off (TTO) are two important techniques for the elicitation of health state utilities [1,2]. They are choice-based techniques and are therefore from an economist's point of view preferred to choiceless methods such as the visual analogue scale [3]. The SG and the TTO differ in two important respects. First, the SG is framed in terms of risk, whereas the TTO is riskless. The risk in the SG is commonly analyzed by assuming expected utility, i.e., linearity in probability. The second difference is that the TTO assumes linear utility for duration. The SG imposes no restriction on the utility for duration.

Empirical studies have shown that the SG and the TTO yield systematically different utilities [4–9]. The common pattern is that SG utilities exceed TTO utilities. The traditional explanation for this pattern is that people have concave utility for duration [3]. If utility for duration is concave then TTO utilities are biased downwards [10,11].

This paper argues that the traditional explanation is not complete because it is based on expected utility. Under expected utility, the only reason why SG and TTO utilities can differ is indeed utility curvature. However, empirical evidence abounds that expected utility does not describe individual preferences well. Indirect violations of expected utility were observed for health outcomes by Llewellyn-Thomas *et al*. [12], Rutten-van Mölken *et al*. [13], and Bleichrodt [14]. Three main reasons

*Correspondence to: iMTA/iBMG, Erasmus University, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. E-mail: bleichrodt@bmg.eur.nl; www.bmg.eur.nl/personal/bleichrodt

Received 12 December 2000
Accepted 17 August 2001

why people deviate from expected utility are probability weighting, loss aversion, and scale compatibility. Wakker and Stiggelbout [15] analyzed the impact of probability weighting on SG utilities. In their discussion, Jansen *et al.* [16] mentioned the possible impact of probability weighting, loss aversion, and scale compatibility on SG and TTO utilities. This paper extends these previous studies by offering a detailed analysis of the effect of probability weighting, loss aversion, and scale compatibility on SG and TTO utilities. This analysis can explain the available empirical evidence on SG and TTO utilities and their (in) consistency with individual preferences for health.

This paper will show that the SG is biased upwards by probability weighting (in most cases) and by loss aversion. The effect of scale compatibility on SG utilities is ambiguous. TTO utilities are biased upwards by loss aversion and scale compatibility. Probability weighting does not affect the TTO utilities because the TTO elicitation involves no risk. The assumption of linear utility for duration leads to a downward bias in the TTO utilities. Hence, the joint effect of utility curvature, probability weighting, loss aversion, and scale compatibility is an upward bias in SG utilities and the simultaneous existence of upward and downward biases in TTO utilities. This can explain why SG utilities are systematically higher than TTO utilities.

Probability weighting, loss aversion, and scale compatibility are biases in individual preferences that should be avoided in prescriptive decision analyses. Prescribing which medical program should be implemented is a normative exercise and should therefore be based on a normative theory of decision making. The existing nonexpected utility theories are primarily intended as descriptive theories and none of these theories are currently considered as a viable normative alternative for expected utility. Hence, the deviations from expected utility modeled by these theories are undesirable from a prescriptive point of view and should be avoided in prescriptive analyses. Some authors have argued that because health care resource allocation decisions are prescriptive, and expected utility is still the dominant normative theory, and since the SG can elicit expected utilities, the SG should be used in health utility measurement (e.g. [1]). The elicitation of the utilities to be used in a prescriptive analysis, however, is a descriptive activity and will be susceptible to the biases that cause violations of expected utility. The use of biased utilities will lead to biased resource allocation decisions and the elicitation of utilities should therefore be based on the best descriptive techniques available. That is, techniques in which the joint impact of probability weighting, loss aversion, and scale compatibility is minimized.

Several studies presented empirical evidence that TTO utilities better reflect individual preferences for health than SG utilities [3,8,17,18]. This finding cannot be explained if the difference between TTO utilities and SG utilities is only due to curvature of the utility function over duration, as the traditional explanation posits. Because the TTO imposes a restriction on the utility function for duration, whereas the SG leaves the utility function for duration unrestricted, SG utilities can never be less consistent with individual preferences for health than TTO utilities. The explanation suggested in this paper, however, can explain the tentative finding that TTO utilities are more consistent with individual preferences than SG utilities. The higher consistency follows if the upward and downward biases in the TTO approximately cancel.

The following sections introduce notation and definitions and review utility curvature, the traditional explanation for the difference between SG and TTO utilities. Probability weighting, loss aversion, and scale compatibility are described in the subsequent sections along with explanation of the effects of these three factors on SG and TTO utilities. The conclusion is given in the final section.

Preliminaries

We consider chronic health states. Hence, the outcomes can be described as pairs (Q, T) where Q denotes health status and T duration in years. For example, (Asthma, 10) stands for living for 10 years with asthma. A two-outcome lottery giving (Q_1, T_1) with probability p and (Q_2, T_2) with probability $1-p$ is denoted $((Q_1, T_1),\ p;\ (Q_2, T_2))$. Let $\succcurlyeq$ denote the preference relation 'at least as preferred as' defined over lotteries. By setting $p=1$ in $((Q_1, T_1),\ p;\ (Q_2, T_2))$, $\succcurlyeq$ defines a preference relation over outcomes. Strict preference is denoted by $\succ$ and indifference by $\sim$. Throughout the paper, we assume that lotteries are *rank-ordered*, i.e., the notation $((Q_1, T_1),\ p;\ (Q_2, T_2))$ implies that $(Q_1, T_1) \succcurlyeq (Q_2, T_2)$.

Expected utility holds if the utility of the lottery $((Q_1, T_1),\ p;\ (Q_2, T_2))$ can be written as $pU(Q_1, T_1) + (1-p)U(Q_2, T_2)$. U is a real-valued utility function over outcomes that is unique up to scale and location. An intuitive axiomatization of expected utility is provided by Jensen [19].

Throughout the paper, it is assumed that $U(Q, T) = H(Q)G(T)$ where H and G are utility functions over health status and duration, respectively. Axiomatizations of this decomposition have been given by Miyamoto and Eraker [20] and by Miyamoto *et al*. [21]. Empirical tests provided support for the conditions on which the decomposition depends [20,22]. It is further assumed that more years in full health are always preferred to less.

The standard gamble asks for the probability p that leads to indifference between the outcome (Q_1, T) and the lottery $((Q_2, T),\ p;\ (Q_3, T))$ where $(Q_2, T) \succcurlyeq (Q_1, T) \succcurlyeq (Q_3, T)$. Q_2 is commonly set equal to full health, denoted FH, and (Q_3, T) to death, formally defined as $(Q, 0)$ where Q can be any health state. This convention is followed throughout the paper. We set $H(\mathrm{FH}) = 1$ and $U(\mathrm{death}) = 0$, which is allowed by the uniqueness properties of U. Evaluating the indifference $(Q_1, T) \sim ((\mathrm{FH}, T),\ p;\ \mathrm{death})$ by expected utility yields

$$H(Q_1)G(T) = pH(FH)G(T) + (1-p)0 = pG(T) \tag{1}$$

and thus

$$H(Q_1) = p \tag{2}$$

which is the commonly used formula for the SG utility of health state Q_1 [1].

The time trade-off asks for the duration T_2 that yields indifference between (Q_1, T_1) and (FH, T_2). Utility is *linear in duration* if $U(Q, T) = H(Q_1)\ T$. If utility is linear in duration then the TTO indifference gives

$$H(Q_1)T_1 = H(FH)T_2 = T_2 \tag{3}$$

Hence,

$$H(Q_1) = \frac{T_2}{T_1} \tag{4}$$

which is the commonly used formula for the TTO utility of health state Q_1 [1].

We have used the same utility function $U(Q, T)$ in the evaluation of the SG and of the TTO. Several authors have argued that the utility (or value) function used in a riskless decision context, such as the TTO, differs from the utility function used in a risky decision context, such as the SG [23–25]. The distinction between risky and riskless utility functions was motivated by the fact that diminishing marginal utility and risk attitude cannot be separated under expected utility. The distinction between risky and riskless utility functions may be less relevant under nonexpected utility, where not all risk attitude is reflected in the utility function and people can be risk averse even when the riskless and the risky utility function are identical. Wakker [26] argues in favor of a unified concept of utility that is applicable both in riskless and in risky decisions. Empirical support for such a unified concept of utility is given by Stalmeier and Bezembinder [27] and Abdellaoui *et al*. [28]. Richardson [29] and Dolan [3] present further arguments that the SG and the TTO measure the same underlying utility function.

Utility curvature

The previous section showed that the SG imposes no restriction on the utility function for duration whereas the TTO assumes that utility is linear in duration, i.e., $G(T) = T$. The TTO utility is therefore biased if $G(T)$ is nonlinear. Then the utility of Q_1 is equal to $G(T_2)/G(T_1)$ and not to T_2/T_1 as assumed in the TTO. The following proposition summarizes the consequences of utility curvature for the TTO utilities.

Proposition 1
(a) if G is concave then the TTO utilities are biased downwards,
(b) if G is convex then the TTO utilities are biased upwards.

Because there is no risk in the TTO, Proposition 1 is valid both under expected utility and under nonexpected utility. Figure 1 shows graphically that concave utility leads to a downward bias in the TTO utilities. In the figure G is scaled such that $G(T_1) = T_1$. This is allowed by the uniqueness properties of U. Since $G(T_2) > T_2$ by the concavity of G, it follows that $G(T_2)/G(T_1) > T_2/T_1$. Thus,

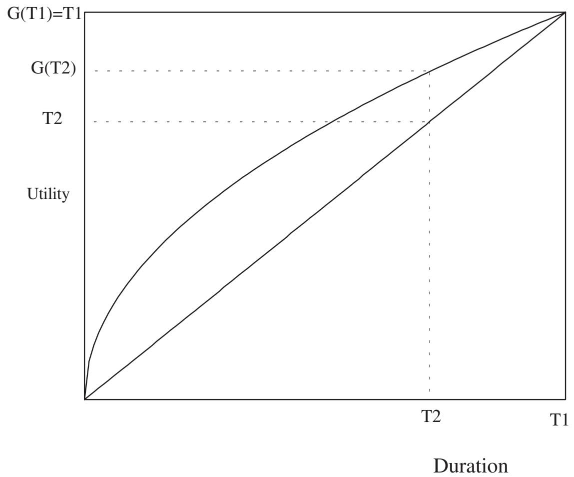

Figure 1. Concavity of the utility function implies that the TTO utilities are biased upwards

the true utility of health state Q_1, $G(T_2)/G(T_1)$, exceeds the TTO utility, T_2/T_1, or, equivalently, the TTO utility is biased downwards.

Johannesson *et al.* [30] and Dolan and Jones-Lee [31] previously argued that the TTO is biased downwards if individuals have positive time preference. Because positive time preference is implied by a concave utility function over duration, Proposition 1 contains their analysis as a special case.

Empirical evidence shows that G is concave [7,32–37]. Hence, these empirical findings suggest by application of Proposition 1 that the TTO utilities are biased downwards. Because the SG imposes no restrictions on G, the existence of utility curvature does not lead to a bias in the SG utilities.

Probability weighting

An important reason why people deviate from expected utility is that they do not evaluate probabilities linearly, as expected utility assumes, but weight probabilities. Evidence of probability weighting is well-documented, both for monetary outcomes [38–44] and for health outcomes [27,37,45]. A formal theory of probability weighting is rank-dependent utility (RDU) theory [46], currently the most influential descriptive theory of decision under risk. Under RDU, the lottery $((Q_1, T), p; (Q_2, T))$ is evaluated as

$$w(p)U(Q_1, T) + (1 - w(p))U(Q_2, T) \tag{5}$$

where w is a probability weighting function with $w(0) = 0$ and $w(1) = 1$. The probability weighting function is strictly increasing, i.e. for all probabilities p and q, $w(p) > w(q)$ if $p > q$. U is, as before, a real-valued utility function, unique up to scale and location.

The TTO utility is elicited under certainty. Hence, no probabilities are involved and probability weighting does not affect the TTO utilities. Probabilities feature in the SG and, therefore, probability weighting affects the SG utilities. If the SG indifference $(Q_1, T) \sim ((\mathrm{FH}, T), p; \text{death})$ is evaluated by Equation (5) using the scaling $H(\mathrm{FH}) = 1$ and $U(\text{death}) = 0$, we obtain

$$H(Q_1) = w(p) \tag{6}$$

which is the utility of health state Q_1 under RDU.

An example may illustrate how probability weighting affects the SG utilities. Suppose an individual indicates that he is indifferent between (Q_1, T) and $((\mathrm{FH}, T), 0.70; \text{death})$. Then the SG utility of health state Q_1 is equal to 0.70 by Equation (2). Suppose the individual weights probabilities. More specifically, suppose he underweights probabilities, i.e. for all p, $w(p) < p$. Figure 2 illustrates such a probability weighting function. Then it follows immediately from Equation (6)

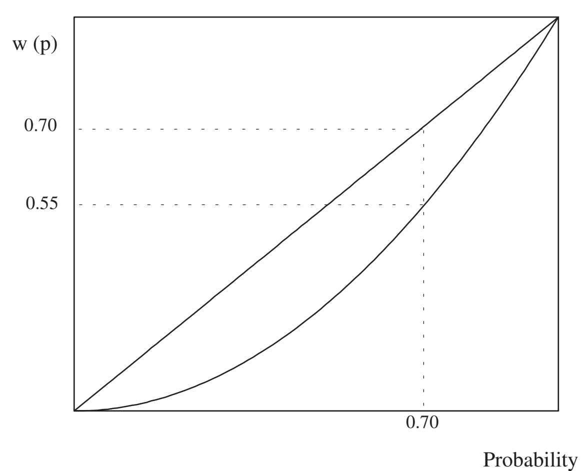

Figure 2. Underweighting of probabilities implies an upward bias in the SG utilities

that the individual's true utility of Q_1 is less than 0.70. In this case, the SG overestimates the utility of Q_1.

Let us now summarize the impact of probability weighting on the SG utility.

Proposition 2. If RDU holds then
(a) if $w(p) > p$ for all p, i.e. the individual overweights all probabilities, then the SG utilities are biased downwards,
(b) if $w(p) < p$ for all p, i.e. the individual underweights all probabilities, then the SG utilities are biased upwards.

Empirical evidence shows that the probability weighting function is typically inverse S-shaped [37,40–45,47,48]. Figure 3 illustrates an inverse S-shaped probability weighting function. The inverse S-shaped probability weighting function overweights small probabilities and underweights larger probabilities. The point where the function changes from overweighting probabilities to underweighting probabilities lies approximately at 0.35. In general, the probabilities that are reported in SG elicitations are well above 0.35 [18,49,50]. Hence, the inverse S-shaped probability weighting function implies in combination with Proposition 2 that the SG utilities are generally biased upwards by probability weighting.

Loss aversion

Kahneman and Tversky [39] argued that an important reason why people deviate from expected utility is that they evaluate outcomes as gains and losses relative to a reference point. Moreover, people are *loss averse*, i.e., they are more sensitive to losses than to gains. Theories of reference dependence and loss aversion have been described by Tversky and Kahneman [51] for certainty and by Tversky and Kahneman [40] for risk. Many empirical papers confirmed the importance of loss aversion and related phenomena as the endowment effect and status quo bias [51–54]. Kahneman and Tversky [39], McNeil *et al.* [55], Stalmeier and Bezembinder [27] and Bleichrodt and Pinto [56] present evidence of the impact of loss aversion on medical trade-offs.

In the TTO, an individual is asked to state the number of years T_2 in full health that makes him indifferent between (Q_1, T_1) and (FH, T_2). In terms of a reference-dependent theory, this means that the individual's reference point is (Q_1, T_1) and that he is 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. Let T_2' be the individual's answer if he is equally sensitive to gains and losses. That is, T_2' is the individual's answer in the absence of loss aversion. This answer implies that the utility of the gain $(\text{FH}-Q_1)$ exactly offsets the utility of the loss $(T_2'-T_1)$. If the individual is loss averse the loss $(T_2'-T_1)$ gets more weight than the gain $(\text{FH}-Q_1)$ and therefore the individual will strictly prefer (Q_1, T_1) to (FH, T_2'). To restore indifference, T_2' has to increase, say to T_2''. Hence, the TTO utility under loss aversion, T_2''/T_1, exceeds the TTO utility in the absence of loss aversion, T_2'/T_1, and thus loss aversion leads to an upward bias in the TTO utility.

Figure 4 gives a graphical illustration of the above argument. Loss aversion makes people more reluctant to give up life-years in the TTO because they are more sensitive to losses than to gains. Hence, loss aversion implies that the indifference curves originating from the reference point (Q_1, T_1) become steeper. Consequently, the value of T_2 for which indifference holds between (Q_1, T_1) and (FH, T_2) increases and therefore the TTO utility increases.

It is interesting to note that if we had asked the TTO question in reversed order, i.e., if we had asked how many years in Q_1 makes the individual indifferent to T_2 years in full health, then loss aversion would have predicted a downward bias in the TTO utility. This follows because in this case the individual trades off the gain in duration

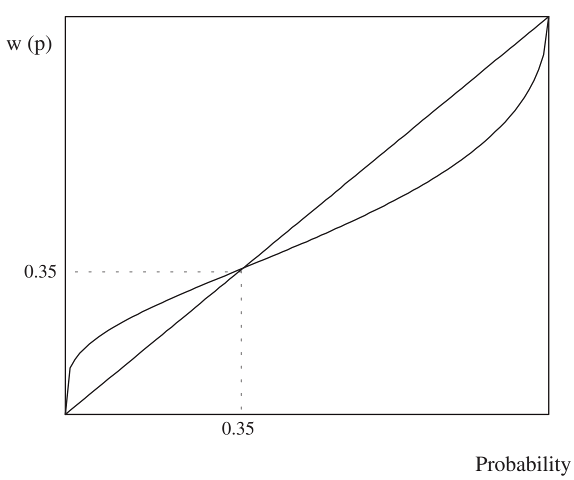

Figure 3. An inverse S-shaped probability weighting function

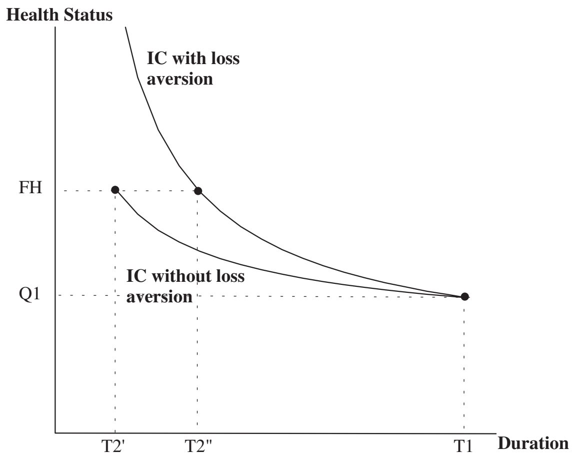

Figure 4. Loss aversion leads to an upward bias in the TTO utilities

from T_1 to T_2 against the loss in health status from FH to Q_1.

Let us now turn to the effect of loss aversion on the SG utility. First, we have to find the location of the individual's reference point. It is hypothesized that the individual will take the outcome that is received with certainty, (Q_1, T_1), as his reference point in the evaluation of the SG question. Hershey and Schoemaker [57] and Bleichrodt *et al.* [58] present evidence that supports this hypothesis. If (Q_1, T_1) is his reference point, the individual trades off the gain from (Q_1, T_1) to (FH, T_1) with probability p against the loss from (Q_1, T_1) to death with probability $1-p$ when he answers the SG. Let p' be the individual's response to the SG question in the absence of loss aversion. That is, obtaining the gain from (Q_1, T_1) to (FH, T_1) with probability p' is just sufficient to offset the loss from (Q_1, T_1) to death with probability $1-p'$ if the individual is equally sensitive to gains and losses. Hence, if the individual is loss averse and therefore more sensitive to losses than to gains then obtaining the gain from (Q_1, T_1) to (FH, T_1) with probability p' is not sufficient to offset the loss from (Q_1, T_1) to death with probability $1-p'$. Consequently, (Q_1, T_1) is strictly preferred to $((\mathrm{FH}, T_1), p'; \text{death})$ if the individual is loss averse and to restore indifference p' has to increase, say to p''. Because $p'' > p'$, the SG utility in the presence of loss aversion exceeds the SG utility in the absence of loss aversion. Hence, loss aversion leads to an upward bias in the SG utility.

Scale compatibility

Scale compatibility means that an individual assigns more weight to an attribute the higher its compatibility with the response scale used. A theory of scale compatibility is described by Tversky *et al.* [59]. Delquié [60,61] presents extensive empirical evidence of the impact of scale compatibility on individual preferences. Bleichrodt and Pinto [56] find scale compatibility in medical trade-offs.

In the TTO, the individual is asked how many years in full health are equivalent to (Q_1, T_1), and so the response scale is duration. Scale compatibility then implies that the individual will give more weight to duration than to health status in answering the TTO question. Let T_2' be the response of an individual who exhibits no scale compatibility. This response implies that being in full health instead of in Q_1 exactly compensates for living only T_2' years instead of T_1 years when the individual's preferences are not affected by the response scale used. Suppose now that the individual exhibits scale compatibility, i.e., he gives more weight to duration than to health status in responding to the TTO questions. Then (Q_1, T_1) will be strictly preferred to (FH, T_2') because $T_1 > T_2'$ and the individual gives more weight to duration than to health status compared to the case in which there is no impact of scale compatibility. Hence, being in full health rather than in Q_1 can no longer compensate for living for T_2' years instead of T_1 years. To restore indifference, the number of years in full health has to increase, say to T_2''. Because $T_2'' > T_2'$, the TTO utility T_2''/T_1 of an individual who exhibits scale compatibility exceeds the TTO utility T_2'/T_1 of an individual whose preferences are insensitive to the response scale used.

Figure 5 illustrates the effect of scale compatibility on the TTO utilities. Because duration is used as the response scale, an individual who exhibits scale compatibility will be less willing to give up life-years for a given improvement in health status. Hence, the indifference curves originating from the point (Q_1, T_1) become steeper due to scale compatibility and the value of T_2 for which indifference holds increases. A similar effect was caused by loss aversion. The cause of the upward bias in the TTO utilities is different, however. It can be seen that loss aversion and scale compatibility can have different implications

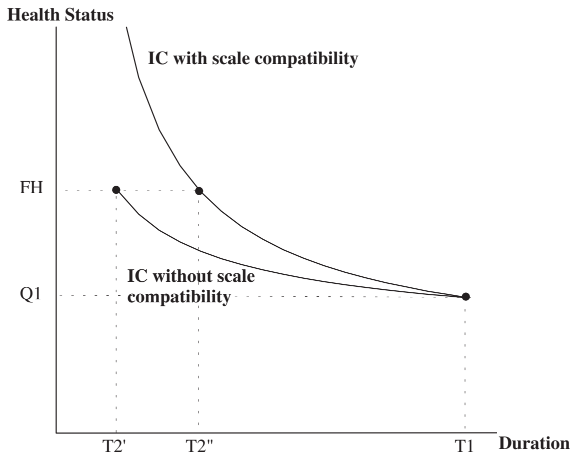

Figure 5. Scale compatibility leads to an upward bias in the TTO utilities

by considering their effects on the reversed TTO question as to how many years in Q_1 are equivalent to T_2 years in FH. Then loss aversion predicts a downward bias in the TTO utility, but scale compatibility an upward bias.

In the SG, probability is used as the response scale. Therefore, scale compatibility predicts that an individual will focus on probability in the evaluation of the SG question. Three probabilities are involved in the SG question: a probability one of the outcome (Q_1, T), a probability p of the outcome (FH, T) and a probability $(1-p)$ of the outcome death. Focusing on one, the probability of (Q_1, T_1) means that (Q_1, T_1) gets extra decision weight and becomes more attractive compared to the treatment option $((\mathrm{FH}, T), p; \text{death})$. To restore indifference, the treatment option must be made more attractive, which is achieved by increasing p. Hence, focusing on the probability of (Q_1, T_1) leads to an increase in the reported indifference probability in the SG question and thus to an upward bias in the SG utility.

If the individual focuses on p, the probability of the good outcome (FH, T_1) in the treatment option, then the treatment option becomes more attractive compared to the certain outcome (Q_1, T_1). To restore indifference, the treatment option must be made less attractive which is achieved by decreasing p. Hence, focusing on the probability p leads to a downward bias in the SG utility.

If the individual focuses on $1-p$, the probability of the bad outcome death in the treatment option, then the treatment option becomes less attractive compared to the certain outcome (Q_1, T_1). To restore indifference, the treatment option must be made more attractive which is achieved by increasing p. Hence, focusing on the probability $1-p$ leads to an upward bias in the SG utility.

We thus observe that scale compatibility leads both to downward and to upward biases in the SG utilities. Scale compatibility does not predict which bias is more pronounced. Therefore, the overall effect of scale compatibility on the SG utilities is ambiguous. Let us finally note that the extra decision weight given to the probabilities in the SG as a result of scale compatibility is different from the type of probability weighting previously discussed. Under scale compatibility, the overweighting of probabilities follows from the use of probability as the response scale. It will disappear if a different response scale is used. Probability weighting as previously discussed follows because people do not evaluate probabilities linearly and will affect the SG utilities regardless of whether probability is used as the response scale.

Conclusion

Table 1 summarizes the analysis of this paper. The table shows that the SG utilities will generally overestimate the utility of a health state. The TTO can both under- and overestimate the utility of a

Table 1. Overview of the biases in the SG and the TTO utilities

Effect	Bias in SG utility	Bias in TTO utility
Utility curvature	None	Downward
Probability weighting	Generally upward	None
Loss aversion	Upward	Upward
Scale compatibility	Ambiguous	Upward

health state depending on the sizes of the different biases.

The conclusions about the impact of probability weighting on the SG utilities depend on the descriptive validity of RDU. The only lotteries that are evaluated in the SG are two-outcome lotteries. Empirical evidence on the descriptive validity of RDU for two-outcome lotteries is generally favorable [62].

The analysis of this paper can explain why SG utilities are higher than TTO utilities. The analysis does not imply necessarily that TTO utilities are closer to the 'true' individual preferences than SG utilities. Whether the TTO is more consistent with individual preferences than the SG depends on the extent to which the different biases in the TTO cancel. As noted in the introduction, there is some evidence that the TTO is more consistent with individual preferences than the SG. Moreover, Bleichrodt and Johannesson [8] found that the consistency of the TTO with individual preferences is maximal for a discount rate approximately equal to zero, i.e., when the utility function for duration is linear. Their finding suggests that the downward bias in the TTO utility caused by utility curvature approximately offsets the upward bias caused by loss aversion and scale compatibility. Obviously, more research must be performed before this tentative conclusion can be firmly established.

It is interesting to note that the assumption of linear utility, which is often believed to be a weakness of the TTO, is crucial in the explanation for why the TTO can be more consistent with individual preferences than the SG. Without this assumption the TTO would also be biased upwards due to loss aversion and scale compatibility. This observation implies that proposals to adjust TTO measurements for utility curvature [30,31] may actually decrease the consistency of the TTO with individual preferences.

The conclusions of this paper are based on aggregate findings. Empirical research shows that much variation exists at the individual level and the direction of the biases in the SG and the TTO may differ from the predictions of this paper at the individual level. The existence of biases in SG and TTO utilities, the direction of which cannot always be predicted, highlights that the major challenge for health utility measurement is to develop utility measures and/or utility elicitation procedures that avoid or minimize the impact of biases on health utilities.

Acknowledgements

Jose Luis Pinto, Peter Wakker, and two referees made many helpful comments. Han Bleichrodt's research was made possible by a fellowship from the Royal Netherlands Academy of Arts and Sciences.

References

1. Torrance GW. Measurement of health state utilities for economic appraisal: a review. *J Health Econ* 1986; **5**: 1–30.
2. Drummond MF, O'Brien B, Stoddart GL, Torrance GW. *Methods for the Economic Evaluation of Health Care Programmes*. Oxford University Press: New York, 1997.
3. Dolan P. The measurement of health-related quality of life for use in resource allocation decisions in health care. In *Handbook of Health Economics*, vol. 1B, Culyer AJ, Newhouse JP. (eds) Elsevier Science: Amsterdam, 2000; 1723–1760.
4. Torrance GW. Social preferences for health states: an empirical evaluation of three measurement techniques. *Socio-Econ Planning Sci* 1976; **10**: 129–136.
5. Read JL, Quinn DM, Berrick HV, Fineberg HV, Weinstein ML. Preferences for health outcomes: comparison of assessment methods. *Med Decision Making* 1984; **4**: 315–329.
6. Hornberger JC, Redelmeier DA, Petersen J. Variability among methods to assess patients well-being and consequent effect on a cost-effectiveness analysis. *J Clin Epidemiol* 1992; **45**: 505–512.
7. Stiggelbout AM, Kiebert GM, Kievit J, Leer JWH, Stoter G, de Haes JCJM. Utility assessment in cancer patients: adjustment of time tradeoff scores for the utility of life years and comparison with standard gamble scores. *Med Decision Making* 1994; **14**: 82–90.
8. Bleichrodt H, Johannesson M. Standard gamble, time trade-off, and rating scale: experimental results on the ranking properties of QALYs. *J Health Econ* 1997; **16**: 155–175.
9. Lenert LA, Cher DJ, Goldstein MK, Bergen MR, Garber AM. The effect of search procedures on utility elicitations. *Med Decision Making* 1998; **18**: 76–83.
10. Miyamoto JM, Eraker SA. Parameter estimates for a QALY utility model. *Med Decision Making* 1985; **5**: 191–213.
11. Cher DJ, Miyamoto J, Lenert LA. Incorporating risk attitude into Markov-process decision models. *Med Decision Making* 1997; **17**: 340–350.
12. Llewellyn-Thomas H, Sutherland HJ, Tibshirani R, Ciampi A, Till JE, Boyd NF. The measurement of patients' values in medicine. *Med Decision Making* 1982; **2**: 449–462.

13. Rutten-van Mölken MP, Bakker CH, van Doorslaer EKA, van der Linden S. Methodological issues of patient utility measurement. Experience from two clinical trials. *Med Care* 1995; **33**: 922–937.
14. Bleichrodt H. Probability weighting in choice under risk: an empirical test. *J of Risk Uncertainty* 2001; **23**: 185–198.
15. Wakker PP, Stiggelbout AM. Explaining distortions in utility elicitation through the rank-dependent model for risky choices. *Med Decision Making* 1995; **15**: 180–186.
16. Jansen SJT, Stiggelbout AM, Wakker PP, Vliet Vlieland TPM, Leer J-WH, Nooy MA, Kievit J. Patient expected utilities for cancer treatments: a study on the feasibility of a chained procedure for the standard gamble and the time trade-off. *Med Decision Making* 1998; **18**: 391–399.
17. van Busschbach JJ. *The Validity of Qalys*. Gouda Quint: Arnhem, The Netherlands, 1994 (in Dutch).
18. Dolan P, Gudex C, Kind P, Williams A. Valuing health states: a comparison of methods. *J Health Econ* 1996; **15**: 209–231.
19. Jensen NE. An introduction to Bernoullian utility theory: I. Utility functions. *Scand J Econ* 1967; **69**: 163–183.
20. Miyamoto JM, Eraker SA. A multiplicative model of the utility of survival duration and health quality. *J Exp Psychol: General* 1988; **117**: 3–20.
21. Miyamoto JM, Wakker PP, Bleichrodt H, Peters HJM. The zero-condition: a simplifying assumption in QALY measurement and multiattribute utility. *Manage Sci* 1998; **44**: 839–849.
22. Bleichrodt H, Pinto JL. The validity of QALYs under non-expected utility. *Working Paper Erasmus University*, 2001.
23. Keeney R, Raiffa H. *Decisions with Multiple Objectives*. Wiley: New York, 1976.
24. Gafni A, Birch S, Mehrez A. Economics, health, and health economics: HYEs versus QALYs. *J Health Econ* 1993; **12**: 325–339.
25. Morrison GC. HYE and TTO: what is the difference? *J Health Econ* 1997; **16**: 563–578.
26. Wakker PP. Separating marginal utility and probabilistic risk aversion. *Theory Decision* 1994; **36**: 1–44.
27. Stalmeier PFM, Bezembinder TGG. The discrepancy between risky and riskless utilities: a matter of framing? *Med Decision Making* 1999; **19**: 435–447.
28. Abdellaoui M, Barrios C, Wakker PP. Did vNM resurrect cardinal utility after all? Theoretical and empirical arguments based on non-expected utility. Working Paper, ENS, Cachan, 2000.
29. Richardson J. Cost utility analysis: what should be measured? *Soc Sci Med* 1994; **39**: 7–22.
30. Johannesson M, Pliskin JS, Weinstein MC. A note on QALYs, time tradeoff, and discounting. *Med Decision Making* 1994; **14**: 188–193.
31. Dolan P, Jones-Lee M. The time trade-off: a note on the effect of lifetime reallocation of consumption and discounting. *J Health Econ* 1997; **16**: 731–739.
32. McNeil BJ, Weichselbaum R, Pauker SG. Fallacy of the five-year survival in lung cancer. *N Engl J Med* 1978; **299**: 1397–1401.
33. Verhoef LCG, de Haan AFJ, van Daal WAJ. Risk attitude in gambles with years of life: empirical support for prospect theory. *Med Decision Making* 1994; **14**: 194–200.
34. Wakker PP, Deneffe D. Eliciting von Neumann–Morgenstern utilities when probabilities are distorted or unknown. *Manage Sci* 1996; **42**: 1131–1150.
35. Chapman G. Temporal discounting and utility for health and money. *J Exp Psychol: Learn Mem Cognit* 1996; **22**: 771–791.
36. Cairns JA, van der Pol MM. Saving future lives: a comparison of three discounting models. *Health Econ* 1997; **6**: 341–350.
37. Bleichrodt H, Pinto JL. A parameter-free elicitation of the probability weighting function in medical decision analysis. *Manage Sci* 2000; **46**: 1485–1496.
38. Edwards WA. Subjective probabilities inferred from decisions. *Psychol Rev* 1962; **69**: 109–135.
39. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica* 1979; **47**: 263–291.
40. Tversky A, Kahneman D. Advances in prospect theory: cumulative representation of uncertainty. *J Risk Uncertainty* 1992; **5**: 297–323.
41. Tversky A, Fox C. Weighing risk and uncertainty. *Psychol Rev* 1995; **102**: 269–283.
42. Wu G, Gonzalez R. Curvature of the probability weighting function. *Manage Sci* 1996; **42**: 1676–1690.
43. Gonzalez R, Wu G. On the form of the probability weighting function. *Cognit Psychol* 1999; **38**: 129–166.
44. Abdellaoui M. Parameter-free elicitation of utilities and probability weighting functions. *Manage Sci* 2000; **46**: 1497–1512.
45. Bleichrodt H, van Rijn J, Johannesson M. Probability weighting and utility curvature in QALY based decision making. *J Math Psychol* 1999; **43**: 238–260.
46. Quiggin J. Risk perception and risk aversion among Australian farmers. *Aust J Agric Econ* 1981; **25**: 160–169.
47. Lattimore PM, Baker JR, Witte AD. The Influence of probability on risky choice. *J Econ Behavr Organ* 1992; **17**: 377–400.
48. Camerer CF, Ho T-H. Nonlinear weighting of probabilities and violations of the betweenness axiom. *J Risk Uncertainty* 1994; **8**: 167–196.
49. Lundberg L, Johannesson M, Dag GL, Borquist L. The relationship between health state utilities and

the SF-12 in a general population. *Med Decision Making* 1999; **19**: 128–141.
50. Lenert, LA, Sherbourne CD, Reyna V. Utility elicitation using single-item questions compared with a computerized interview. *Med Decision Making* 2001; **21**: 97-104.
51. Tversky A, Kahneman D. Loss aversion in riskless choice: a reference-dependent model. *Q J Econ* 1991; **56**: 1039–1061.
52. Samuelson W, Zeckhauser R. Status quo bias in decision making. *J Risk Uncertainty* 1988; **1**: 7–59.
53. Kahneman D, Knetsch JL, Thaler RH. Experimental test of the endowment effect and the Coase theorem. *J Political Econ* 1990; **98**: 1325–1348.
54. Bateman I, Munro A, Rhodes BCS, Sugden RA. Test of the theory of reference-dependent preferences. *Q J Econ* 1997; **62**: 479–505.
55. McNeil BJ, Pauker SG, Sox HC, Tversky A. On the elicitation of preferences for alternative therapies. *N Engl J Med* 1982; **306**: 1259–1262.
56. Bleichrodt H, Pinto JL. Loss aversion and scale compatibility in two-attribute trade-offs. *J Math Psychol*, in press.
57. Hershey JC, Schoemaker PJH. Probability versus certainty equivalence methods in utility measurement: are they equivalent? *Manage Sci* 1985; **31**: 1213–1231.
58. Bleichrodt H, Pinto JL, Wakker PP. Using descriptive findings of prospect theory to improve the prescriptive use of expected utility. *Manage Sci* 2001; **47**: 1498–1514.
59. Tversky A, Sattath S, Slovic P. Contingent weighting in judgment and choice. *Psychol Rev* 1988; **95**: 371–384.
60. Delquié P. Inconsistent trade-offs between attributes: new evidence in preference assessment biases. *Manage Sci* 1993; **39**: 1382–1395.
61. Delquié P. 'Bi-Matching': a new preference assessment method to reduce compatibility effects. *Manage Sci* 1997; **43**: 640–658.
62. Luce RD. *Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches*. Lawrence Erlbaum Associates, Inc.: Mahwah, NJ, 2000.

