

# MODELS OF QUALITY-ADJUSTED LIFE YEARS WHEN HEALTH VARIES OVER TIME: SURVEY AND ANALYSIS

Kristian Schultz Hansen

*Department of Epidemiology and Social Medicine, University of Aarhus*

Lars Peter Østerdal

*Department of Economics, University of Copenhagen*

**Abstract.** Quality-adjusted life year (QALY) models are widely used for economic evaluation in the health care sector. In the first part of the paper, we establish an overview of QALY models where health varies over time and provide a theoretical analysis of model identification and parameter estimation from time trade-off (TTO) and standard gamble (SG) scores. We investigate deterministic and probabilistic models and consider five different families of discounting functions in all. The second part of the paper discusses four issues recurrently debated in the literature. This discussion includes questioning the SG method as the gold standard for estimation of the health state index, re-examining the role of the constant-proportional trade-off condition, revisiting the problem of double discounting of QALYs, and suggesting that it is not a matter of choosing between TTO and SG procedures as the combination of these two can be used to disentangle risk aversion from discounting. We find that caution must be taken when drawing conclusions from models with chronic health states to situations where health varies over time. One notable difference is that in the former case, risk aversion may be indistinguishable from discounting.

**Keywords.** Medical decision making; QALY; Standard gamble; Time trade-off; Constant proportional trade-off; Double discounting

## 1. Introduction

Models of quality-adjusted life years (QALYs) are widely used as health-related quality of life measures for economic evaluation in the health care sector. In its simplest form, the QALY approach assumes that the utility of a profile of health states over time may be estimated by adding the values of the constituent health states in the health profile. In more sophisticated models, health events occurring in the future may be discounted or risk adjusted.

In much of the theoretically oriented literature on QALY measurement, attention has been restricted to situations with chronic health states (e.g. Pliskin *et al.*, 1980; Gafni and Torrance, 1984; Miyamoto and Eraker, 1985, 1988, 1989; Gafni, 1994; Johannesson *et al.*, 1994; Bleichrodt *et al.*, 1997; Miyamoto *et al.*, 1998; Miyamoto, 1999; Bleichrodt, 2002; Doctor and Miyamoto, 2003; Østerdal, 2005).

Restricting attention to chronic health states is indeed very powerful. For instance, elegant axiomatic characterizations of QALY models can be established under expected utility or rank-dependent utility assumptions (Bleichrodt *et al.*, 1997; Miyamoto *et al.*, 1998; Miyamoto, 1999), and it is a useful simplification for an axiomatic approach to distributive justice in health care resource allocation (Østerdal, 2005).

For some applications, models where health states are assumed to be chronic may be sufficiently rich. This is the case in situations where we aim to assess the benefit from a health care service that remedies a permanent handicap or from a health care program that with some probability can extend the life for people with a chronic disease. In practice, however, we often need to consider medical conditions where the health state varies over time.

Other papers have addressed health preference measurement in situations where health varies over time (e.g. Lipscomb, 1989; Stiggelbout *et al.*, 1994; Richardson *et al.*, 1996; Cher *et al.*, 1997; Kuppermann *et al.*, 1997; Krabbe and Bonsel, 1998; Treadwell, 1998; MacKeigan *et al.*, 2003; Spencer, 2003). For instance, recent research has investigated the validity of the additive independence assumption of QALY models. This assumption implies for instance that the utility of an individual health event (at a particular point in time) does not depend on any other health event in the health profile (Richardson *et al.*, 1996; Krabbe and Bonsel, 1998; Treadwell, 1998; Spencer, 2003). The present paper is closely related to the contributions mentioned, but our focus is different.

In the first part of the paper (Sections 2 and 3), we present a unified approach to health preference measurement in situations where health states are not necessarily chronic. This serves as an overview of QALY models that makes it easy to see the shared features and differences between various QALY models. We distinguish between deterministic and probabilistic models and consider four different types of discounting (five in all with a special variant discussed in Section 3.5). In each case, we specify the type of data that is in theory sufficient for estimating the parameters in the models. We focus on procedures using either standard gamble (SG) scores, time trade-off (TTO) scores, or a combination of the two, involving only comparisons of chronic health states, types of data that have been collected in many choice experiments (see, e.g., Dolan *et al.*, 1996; Bleichrodt and Johannesson, 1997).

In Section 2, we consider in turn four deterministic models with exponential, power, proportional, and hyperbolic discounting respectively and describe how parameters may be estimated from TTO scores (Sections 2.1–2.4).

In Section 3, we go on to models where uncertainty is explicit and describe how parameters can be estimated with a combination of TTO and SG scores for each of the discounting models (Sections 3.1–3.4). In Section 3.5, we also consider a

fifth type of discounting related to work by Johannesson *et al.* (1994). In this case, we initially restrict attention to chronic health states and then discuss possible generalizations of the model to situations with non-chronic health states.

Knowing the mathematical relationship between TTO and SG scores and between the time horizon and TTO scores, an appropriate discounting model can in principle be identified from plots of the empirical relationship of these. However, it turns out that the identification of the 'true' model may be extremely difficult even if such a true model exists. Some issues related to model identification are discussed in Sections 2.5 and 3.6.

The second part of the paper (Section 4) contains a discussion based on these models and the above-mentioned findings, and we will discuss some conclusions drawn in the existing literature.

First, we discuss the claim that has been made that the SG method is the 'gold standard' for estimating the health state index and argue against its general validity.

Second, we re-examine the role of the constant-proportional trade-off assumption of life years for health state. In contrast to what has been previously argued in the literature, we find that this condition is not required for the SG method and that it does not rule out discounting.

Third, we discuss the so-called 'double discounting of QALYs' problem. This phenomenon refers to a potential bias related to the use of a discounting function in conjunction with a health state index that is already influenced by time preferences. Whereas we acknowledge the importance of correcting for this bias, we find that 'double discounting' is hardly a meaningful terminology in relation to the use of a SG-based health state index. Furthermore, we argue against the validity of attempting to adjust for double discounting by reducing or even eliminating the use of discounting.

Fourth, in contrast to previous literature often interpreting the TTO method as an alternative to the SG method (and vice versa), we argue that it may not be a matter of choosing between TTO and SG procedures. Rather, it may be useful to utilize both procedures to disentangle risk aversion from discounting.

Finally, we notice that results drawn from theoretical analysis of models involving only chronic health states may not necessarily apply to situations where health is allowed to vary over time; hence, restricting attention to chronic health states may blur some important aspects of QALY modeling. One important difference, among others, is that although risk aversion and discounting (impatience) are generally distinct phenomena, these two effects cannot be disentangled in some of the most widely used models involving only chronic health states, a point which, to our knowledge, has received little attention (if any) in the literature.

The concluding Section 5 mentions three important limitations of our analysis and in relation to this speculates on directions for future research.

## 2. QALY as Utility of Deterministic Health Profiles

In this section, we consider deterministic QALY models. Let  $A$  denote a collection of possible *health states* with typical elements  $a$  and  $a'$ . We make

no suppositions on the nature of such health states, but we shall assume that  $A$  contains a health state  $a^*$  called 'perfect health' and another health state  $a^0$  called 'dead.'

Suppose that an individual is born at time zero and lives for  $t \geq 0$  years. The *health profile* of the individual can then be described by a function  $l$  on the non-negative reals taking values in  $A$ , where  $l(s) \neq a^0$  for all  $s < t$  and  $l(s) = a^0$  for all  $s \geq t$ .

Let  $L$  be a family of possible health profiles, and let  $\succsim$  be a preference relation (a complete and transitive binary relation) on  $L$ . A real-valued function  $q$  on  $L$  represents the preference relation  $\succsim$  on health profiles if

$$l \succsim l' \Leftrightarrow q(l) \geq q(l'), \quad (2.1)$$

for all  $l, l' \in L$ . In QALY models, it is assumed that preferences satisfy the additive independence assumption,<sup>1</sup> and with very little loss of generality, we can assume that there exists a real-valued function  $v$  on  $A \times \mathbb{R}$  such that

$$q(l) = \int v[l(s), s] ds. \quad (2.2)$$

Suppose that  $v(a, s) = u(a)\delta(s)$  for some functions  $u$  and  $\delta$ , where  $0 \leq u \leq 1$  and  $0 < \delta \leq 1$ .<sup>3</sup> In addition, assume that  $u(a^0) \leq u(a) \leq u(a^*)$  for all  $a$  and  $u(a^0) < u(a^*)$ . Without loss of generality, we normalize  $u$  such that  $u(a^0) = 0$  and  $u(a^*) = 1$ . The function  $u$  can be called a *health state index*. Then for any life time  $t$  and health state  $a$ , there is a uniquely determined real number  $h(t, a)$  (which lies between 0 and  $t$ ) such that

$$\int_0^{h(t, a)} u(a^*)\delta(s) ds = \int_0^t u(a)\delta(s) ds.$$

In words,  $h(t, a)$  is the number of years at  $a^*$  (perfect health) that gives the same utility as  $t$  years at health state  $a$  and is referred to as the *TTO index* for chronic health states (e.g. Torrance, 1986; Drummond *et al.*, 1997; Zweifel and Breyer, 1997).

We focus on the most commonly used discounting functions; exponential discounting:  $\delta(s) = c^s$ ,  $0 < c < 1$  (Section 2.1), power discounting:  $\delta(s) = s^z$ ,  $-1 < z < 0$  (Section 2.2), proportional discounting:  $\delta(s) = \frac{1}{1+vs}$ ,  $0 < v$  (Section 2.3), and a two-parameter family of discounting functions that we shall refer to as 'hyperbolic discounting':  $\delta(s) = (1 + vs)^{-w/v}$ ,  $0 < v, w$  (Section 2.4).

In Figure 1, an example of each type of discounting is given where parameters have been selected such that the curves intersect at  $s = 10$ .<sup>4</sup>

The plausibility of various structural forms may be assessed by evaluating underlying axioms which in certain combinations give rise to different functional forms of health-related utility (e.g. Miyamoto and Eraker, 1988; Bleichrodt and Johannesson, 2001). However, this line of analysis is beyond the scope of the present paper. In the following, our focus is parameter estimation (utility assessment) and model identification from TTO scores.

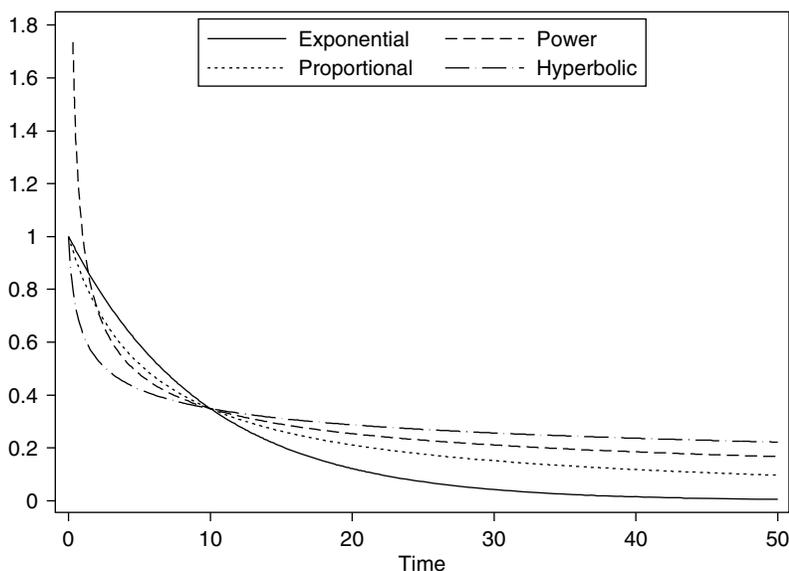


Figure 1. Four types of Discounting.

2.1. Exponential Discounting

With exponential discounting, the QALY model has the form

$$q(t) = \int u(l(s))c^s ds, \tag{2.3}$$

for some health state index  $u$  and  $0 < c < 1$ . The question is how to estimate the health state index  $u$  and the discounting parameter  $c$  from empirical data. For this, we can make use of TTO scores involving only chronic health states. Let  $a$  be a health state such that  $0 < h(t, a) < t$ . We then have the following relation

$$\int_0^{h(t,a)} c^s ds = \int_0^t u(a)c^s ds,$$

where  $u(a)$  and  $c$  are unknowns. This equation reduces to

$$c^{h(t,a)} - 1 = u(a)(c^t - 1). \tag{2.4}$$

For  $0 < t < t'$ , we obtain by (2.4) and substitution

$$(c^{h(t,a)} - 1)(c^{t'} - 1) = (c^{h(t',a)} - 1)(c^t - 1).$$

From two empirical estimates  $\widehat{h}(t, a)$  and  $\widehat{h}(t', a)$ , we may then determine  $c$  numerically. For example, if  $\widehat{h}(1, a) = 0.5$  and  $\widehat{h}(2, a) = 0.9$  then  $\widehat{c} \cong 0.67$ .

It remains to estimate the health state index  $u(a)$ . From (2.4),

$$u(a) = \frac{1 - c^{h(1,a)}}{1 - c},$$

so in our example,  $\hat{u}(a) \cong 0.55$ . The derivation above is similar to those offered by Olsen (1994), Cher *et al.* (1997), and Martin *et al.* (2000), and it was also applied, e.g., in Gyrd-Hansen (2002) and Stavem *et al.* (2002) for elicitation of the discounting parameter.

Solving (2.4) for  $h(t, a)$  gives

$$h(t, a) = \frac{\ln[1 + u(a)(c^t - 1)]}{\ln c}. \quad (2.5)$$

Clearly, we have  $u(a)t > h(t, a)$  for any  $0 < c < 1$  and  $u(a)t \rightarrow h(t, a)$  for  $c \rightarrow 1$ . In fact for  $0 < u < 1$ ,  $h(t, a)$  is strictly concave in  $t$  as

$$h'_t(t, a) = u \frac{c^t}{1 + u(c^t - 1)}$$

is positive and

$$h''_{tt}(t, a) = -uc^t(\ln c) \frac{-1 + u}{(1 + uc^t - u)^2}$$

is negative. Plots of  $h(t, a)$  against  $t$  are depicted in Section 2.5 for selected values of  $u$ .

## 2.2. Power Discounting

An axiomatic characterization of power discounting was provided by Harvey (1986) in an integer-time model. See, e.g., Pliskin *et al.* (1980) and Miyamoto and Eraker (1985) for studies of power discounting in a QALY context and, e.g., Cairns and van der Pol (2000), van der Pol and Cairns (2002), and van der Pol and Roux (2005) for other health-related applications.

With power discounting, the QALY model becomes

$$q(l) = \int u[l(s)]s^z ds, \quad (2.6)$$

where  $-1 < z < 0$ . This model satisfies a generalized form of *constant proportional trade-off* of life years for health state (Pliskin *et al.*, 1980). Given a health profile  $l$  and  $d > 0$ , let  $l_d$  be the health profile obtained from 'stretching out' or 'contracting'  $l$  by a factor  $d$  such that the health state at time  $s$  in  $l$  is shifted to time  $ds$  in  $l_d$ , i.e.  $l_d(s) \equiv l(ds)$ ,  $s \geq 0$ . The ranking of two lives does not depend on the scale of measuring life years:  $l \succsim l'$  if and only if  $l_d \succsim l'_d$  for all  $d > 0$ .

For a health state  $a$ , we have

$$\int_0^{h(t,a)} s^z ds = \int_0^t u(a)s^z ds,$$

which reduces to

$$\frac{1}{z + 1} [h(t, a)]^{z+1} = \frac{1}{z + 1} u(a)t^{z+1},$$

or

$$h(t, a) = u(a)^{\frac{1}{z+1}}t. \tag{2.7}$$

We therefore have  $h(t, a) = h(1, a)t$ , or  $u(a) = [h(1, a)]^{z+1}$ .

As  $h(t, a)$  is linear in  $t$ , it is not possible to extract information on discounting from estimates of  $h(t, a)$  for varying  $t$ . In other words, the parameter  $z$  cannot be derived from TTO scores involving only chronic health states. For the purpose of estimating  $z$ , we can make use of non-chronic health profiles. A possible parameter estimation technique is the following: Let  $a$  be an arbitrary health state for which  $0 < h(1, a) < 1$ , and let  $l$  be the health profile composed of 1 year in health  $a^*$  followed by 1 year in health state  $a$ . In addition, let  $l^w$  be the health profile with  $w$  years in health state  $a$  followed by  $2 - w$  years at health state  $a^*$ . Now, if  $w$  is such that  $l^w \sim l$ , then

$$\int_0^1 s^z ds + \int_1^2 s^z [h(1, a)]^{z+1} ds = \int_0^w s^z [h(1, a)]^{z+1} ds + \int_w^2 s^z ds,$$

which after some calculations reduces to

$$1 + [h(1, a)]^{z+1}(2^{z+1} - 1) = [h(1, a)]^{z+1}w^{z+1} + 2^{z+1} - w^{z+1}, \tag{2.8}$$

where  $z$  is the unknown parameter. If the number  $w$  that satisfies  $l^w \sim l$  and  $h(1, a)$  are determined empirically, the solution to (2.8) can then be found numerically. For example, if  $\hat{h}(1, a) = \frac{1}{2}$  and  $\hat{w} = 0.75$ , then  $\hat{z} \cong -0.16$  and  $\hat{u}(a) \cong 0.56$ .

As it will be clear in the following, the case of power discounting is the only discounting model examined where the  $h(\cdot, a)$  is not strictly concave in  $t$  (see Section 2.5).

### 2.3. Proportional Discounting

Proportional discounting has been used by for instance Herrnstein (1981) and Mazur (1987) and more recently Cairns and van der Pol (2000) and van der Pol and Cairns (2002). See Harvey (1995) for a characterization in a discrete-time model. For the QALY model, we get

$$q(l) = \int u[l(s)](1 + vs)^{-1} ds,$$

where  $v > 0$ . For a health state  $a$ , we have

$$\int_0^{h(t, a)} (1 + vs)^{-1} ds = \int_0^t u(a)(1 + vs)^{-1} ds,$$

which reduces to

$$\ln[vh(t, a) + 1] = u \ln(tv + 1). \quad (2.9)$$

We may elicit  $v$  from two estimates  $h(t, a)$  and  $h(t', a)$ . For  $0 < t < t'$ , we obtain by (2.9) and substitution

$$\frac{\ln[vh(t, a) + 1]}{\ln(tv + 1)} = \frac{\ln[vh(t', a) + 1]}{\ln(t'v + 1)}.$$

For example, if  $\widehat{h}(1, a) = 0.5$  and  $\widehat{h}(2, a) = 0.9$  then by solving the equation numerically we get  $\widehat{v} \cong 1.04$  and then  $\widehat{u}(a) \cong 0.59$ .

Solving (2.9) for  $h(t, a)$  yields

$$h(t, a) = \frac{(tv + 1)^{u(a)} - 1}{v}, \quad (2.10)$$

which is strictly concave in  $t$ .

#### 2.4. Hyperbolic Discounting

As with power and proportional discounting, hyperbolic discounting entails discount rates that decline over time. For an axiomatic study, see Loewenstein and Prelec (1992). See, e.g., Cairns and van der Pol (1997a, 1997b, 2000) and van der Pol and Cairns (2002) for health-related applications. In our QALY context, the full model becomes

$$q(t) = \int u[l(s)](1 + vs)^{-w/v} ds,$$

where  $v, w > 0$ . The special case  $v = w$  is proportional discounting (cf. Section 2.3). Note that if  $v$  tends to zero, the discounting function converges to exponential discounting with discount rate  $w$  [i.e.  $\lim_{v \rightarrow 0} (1 + vs)^{-w/v} = e^{-wt}$ ].

For a health state  $a$ , we have

$$\int_0^{h(t, a)} (1 + vs)^{-w/v} ds = \int_0^t u(a)(1 + vs)^{-w/v} ds,$$

which for  $w/v \neq 1$  after some calculations<sup>5</sup> reduces to

$$[(1 + vh(t, a))^{(-w/v+1)} - 1] = u(a)[(1 + tv)^{(-w/v+1)} - 1]. \quad (2.11)$$

We may elicit  $v$  and  $w$  from three estimates  $h(t, a)$ ,  $h(t', a)$ , and  $h(t'', a)$ . For  $0 < t < t' < t''$ , we obtain by (2.11) and substitution

$$[1 + vh(t, a)]^{(-w/v+1)} - 1 = \frac{[1 + vh(t'', a)]^{(-w/v+1)} - 1}{(1 + vt'')^{(-w/v+1)} - 1} [(1 + tv)^{(-w/v+1)} - 1],$$

and

$$[1 + vh(t', a)]^{(-w/v+1)} - 1 = \frac{[1 + vh(t'', a)]^{(-w/v+1)} - 1}{(1 + vt'')^{(-w/v+1)} - 1} [(1 + t'v)^{(-w/v+1)} - 1].$$

For completeness, it remains to consider the special case  $v = w$ . In principle, if there is no solution to the above equation system for which  $v \neq w$ , we should test the data with the proportional discounting model.

For example, assume that  $\hat{h}(1, a) = 0.40$ ,  $\hat{h}(2, a) = 0.75$  and  $\hat{h}(3, a) = 1.08$ . Solving these equations numerically yields  $\hat{v} \cong 4$  and  $\hat{w} \cong 2$ . From this, we have  $\hat{u}(a) \cong 0.5$ .<sup>6</sup>

For any pair of  $v$  and  $w$ , the TTO index  $h(t, a)$  is strictly concave in  $t$ . Solving (2.11) for  $h(t, a)$  gives

$$h(t, a) = - \frac{1 - \left[ 1 + u(a)(1 + tv)^{\frac{-w+v}{v}} - u(a) \right]^{\frac{v}{-w+v}}}{v} \tag{2.12}$$

We have

$$h'_t(t, a) = \left[ 1 + u(a)(1 + tv)^{\frac{v-w}{v}} - u(a) \right]^{\frac{w}{v-w}} u(a)(1 + tv)^{-\frac{w}{v}},$$

which is positive, and after some calculations, we obtain

$$h''_{tt}(t, a) = \left[ 1 + u(a)(1 + tv)^{\frac{v-w}{v}} - u(a) \right]^{\frac{2w-v}{v-w}} u(a)(1 + tv)^{-\frac{w+v}{v}} w[u(a) - 1],$$

which is negative. In Figures 2–4, the function  $h(t, a)$  for the case of hyperbolic discounting looks almost linear. Indeed, if  $w < v$  (as we have assumed in

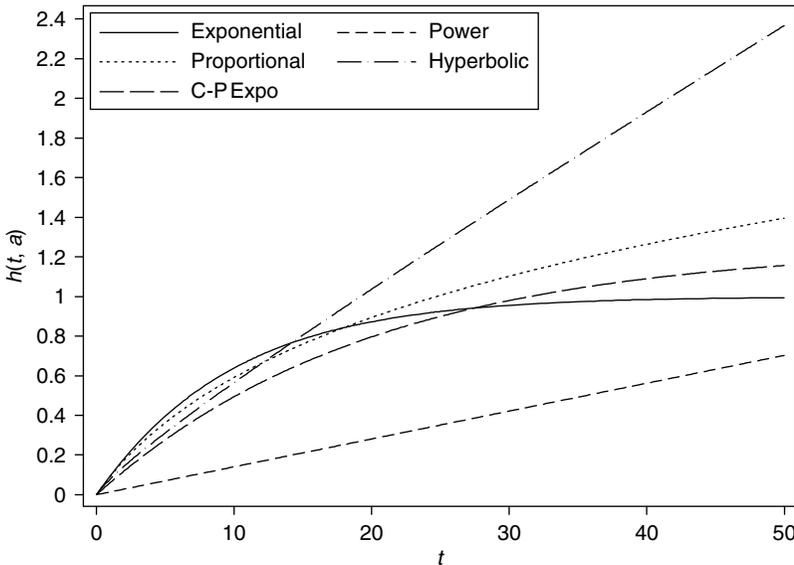


Figure 2. Plot of  $h(t, a)$  against  $t$ ,  $u(a) = 0.1$ .

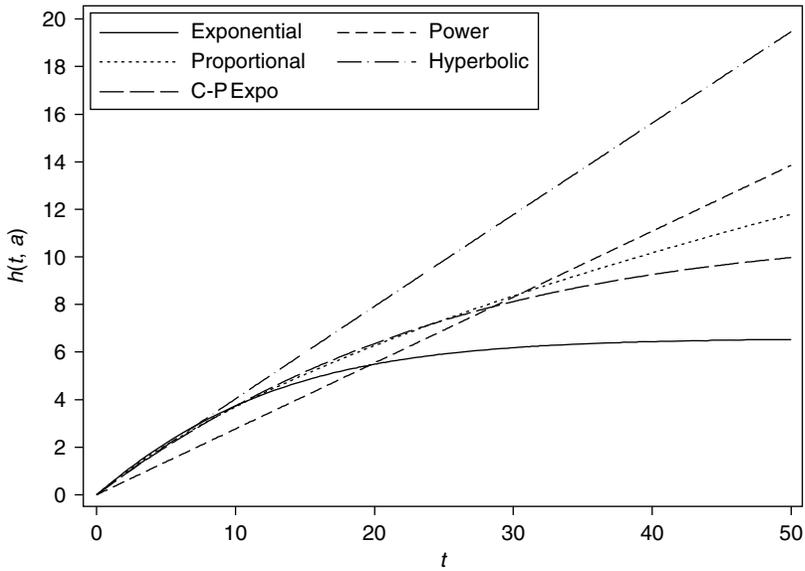


Figure 3. Plot of  $h(t, a)$  against  $t$ ,  $u(a) = 0.5$ .

Figures 2–4),  $h$  is nearly linear in  $t$ . The second order derivatives are small and converging to zero when,  $t$  tends to infinity.<sup>7</sup> If  $w > v$ ,  $h(t, a)$  has an upper bound, however, convergence may be extremely slow.<sup>8</sup>

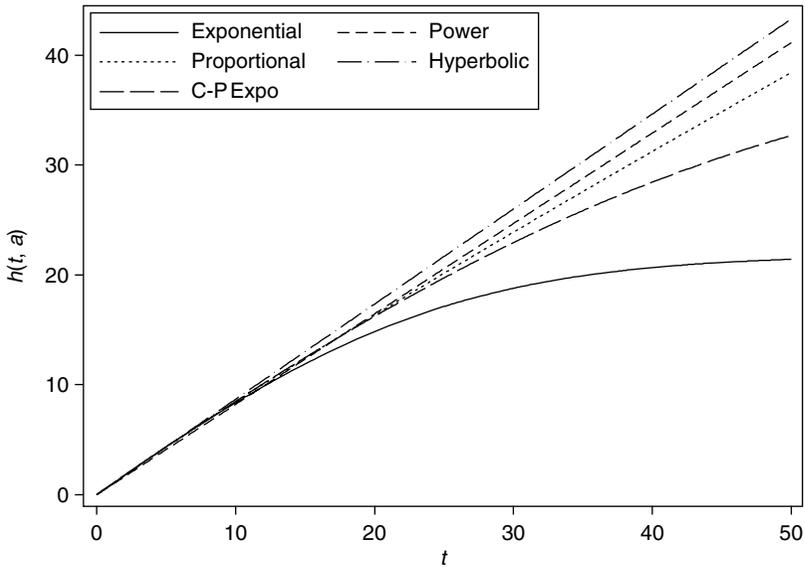


Figure 4. Plot of  $h(t, a)$  against  $t$ ,  $u(a) = 0.9$ .

2.5. Identification of an Appropriate Discounting Model

As we have seen, if the specific type of discounting model in the health domain for an individual is known, the size of the relevant parameter(s) may in principle be identified with little extra effort. In most situations, however, we do not know beforehand which type of discounting function is relevant for a given individual. The most appropriate discounting model for an individual must therefore be identified from experimental data involving a sufficient number of TTO scores for econometric analysis.

A possible procedure could be as indicated in the following. This procedure utilizes the fact that the discounting functions differ qualitatively and will influence the appearance of the function  $h(t, a)$  as seen from equations (2.5), (2.7), (2.10), and (2.12). For a given health state  $a$ , the health index value  $u(a)$  and the TTO index  $h(t, a)$  are a function of the life time  $t$ . The curvature of the TTO index  $h(t, a)$  will thus depend on the discounting model, the size of the discounting parameter(s) and the health state index  $u(a)$ . Examples of the TTO index  $h(t, a)$  have been displayed in the  $(t, h)$  space for specific values of the relevant discounting parameter(s) and  $u(a)$  in Figures 2–4. Keeping this in mind, we can collect TTO scores in a repeated fashion for the same health state but for different remaining life spans  $t$ . (Note that this is different from typical TTO exercises where participants are asked to give scores for a range of health states and a fixed remaining life span.) This will result in a set of  $(t, h)$  pairs for a given health state. Plotting these pairs in a  $(t, h)$ -diagram, an appropriate discounting model may emerge. If a plot of empirical estimates  $\hat{h}(t, a)$  for different values of  $t$  gives approximately a straight line, it may indicate that power or hyperbolic discounting models are well suited. If  $\hat{h}(t, a)$  clearly shows a non-linear relationship, we may wish to select exponential, proportional, or hyperbolic discounting. It is however more likely that the identification of the most appropriate discounting model with associated parameters and the health state index  $u(a)$  for health state  $a$  will require econometric techniques. In other words, the observed  $(t, h)$  values together with an appropriate econometric technique should be used both to select the most suitable discounting model and the parameters of the model.

In Figures 2–4, we use the same parameters as in the examples in Figure 1. For later reference, we also illustrate so-called constant-proportional risk posture over exponentially discounted life years ('C-P Expo' in the figures) which will be discussed in Section 3.5.<sup>9</sup> We consider three cases:  $u = 0.1$  (Figure 2),  $u = 0.5$  (Figure 3), and  $u = 0.9$  (Figure 4).

It is also worth noticing that concavity of  $h(\cdot, a)$  is generally not implied by a positive discounting rate. For example, the discounting function

$$\delta(s) = \begin{cases} s^{-0.5}, & 0 \leq s \leq 1 \\ s^{-0.25}, & 1 < s \end{cases}$$

is continuous and strictly decreasing in  $s$ , but the associated TTO index  $h(\cdot, a)$  is not concave in  $t$ . Concavity of  $h(\cdot, a)$  is therefore a special property shared by all of the discounting functions studied in this paper. Unfortunately, this shared

feature makes it more difficult to identify a model from empirical TTO scores. Hence, if empirical TTO scores seem to follow a strictly *convex* curve, neither of the standard families of discounting functions are likely to provide a good fit to the data, and some non-standard discounting function could be needed.

### 3. QALY as Expected Utility of Health Profiles

In order to deal explicitly with uncertainty, we describe a *health lottery* by a function  $p$  defined on the family of health profiles  $L$  where  $p(l) > 0$  for a finite number of health profiles  $l$ ,  $p(l) = 0$  otherwise, and where all probabilities sum to unity. Let  $P$  denote the set of all such health lotteries on  $L$ .

We now extend the preference relation  $\succsim$  on  $L$  to a preference relation  $\succsim_P$  on  $P$  which represents preferences for lotteries over health profiles.

Let  $\text{supp}(p) = \{l \in L \mid p(l) \neq 0\}$  denote the *support* of a lottery  $p$ . Under expected utility assumptions (see Kreps, 1988, Theorem 5.15), there exists a real-valued function  $q$  on  $L$  such that for any  $p, p' \in P$ ,

$$p \succsim_P p' \Leftrightarrow Q(p) \geq Q(p'), \quad (3.1)$$

where

$$Q(p) = \sum_{l \in \text{supp}(p)} p(l)q(l), \quad (3.2)$$

and where  $q$  is unique up to a positive affine transformation. In QALY models, it is assumed that there is a function  $v$  such that (3.2) holds with  $q(l) = \int v[l(s), s]ds$  (or existence of such functions is derived from assumptions on  $\succsim_P$ ) which gives

$$Q(p) = \sum_{l \in \text{supp}(p)} p(l) \int v[l(s), s]ds. \quad (3.3)$$

In Sections 3.1–3.4, we assume that preferences over health lotteries can be represented by a function  $Q$  of the form (3.3). This model can be interpreted as one with ‘risk neutrality over QALYs’ (see Section 3.5 and Section 4 for a discussion of this assumption). We have risk aversion over QALYs if the expected utility increases from replacing a lottery over QALYs with its expectation, i.e., there is a function  $v$  and a strictly concave function  $f$  such that (3.2) holds with

$$q(l) = f\left(\int v[l(s), s]ds\right). \quad (3.4)$$

Let  $p[t, a, \pi]$  denote a health lottery where, with probability  $\pi$ , health state  $a$  is experienced for  $t$  years followed by death and, with probability  $1 - \pi$ , immediate death occurs. A *SG index* for chronic health states (e.g. Torrance, 1986; Drummond *et al.*, 1997; Zweifel and Breyer, 1997) is then a function  $g$  on  $\mathbb{R}_+ \times A$  taking values in the unit interval, such that  $p[t, a^*, g(t, a)] \sim_P p[t, a, 1]$  for all  $a$  and  $t$ .

Suppose, as in Section 2, that there is  $u$  and  $\delta$  such that  $v(a, s) = u(a)\delta(s)$ , where  $0 < \delta \leq 1$  and  $0 = u(a^0) \leq u(a) \leq u(a^*) = 1$ . In this case, it is easy to verify that  $g$  is well defined and uniquely determined (i.e., we can talk about *the* SG index). As

$$Q(p[t, a^*, g]) = g \int_0^t \delta(s) ds$$

and

$$Q(p[t, a, 1]) = \int_0^t u(a)\delta(s) ds = u(a) \int_0^t \delta(s) ds$$

we have  $g(t, a) = u(a)$  for all  $t$ . In words, under the conditions outlined above, the SG estimate does not depend on the time horizon and is equal to the health state index. In the following, we therefore leave out time as argument in the SG function and write  $g(a) \equiv g(1, a)$ .

In this section, our interest is again the following discounting families: exponential discounting (Section 3.1), power discounting (Section 3.2), proportional discounting (Section 3.3), and hyperbolic discounting (Section 3.4). In all four cases, the parameters can be estimated as outlined in Section 2. In the following, we consider instead procedures using a combination of TTO and SG scores.

In addition, we consider a specific form of discounting derived from an assumption of constant-proportional risk posture over exponentially discounted life years (Section 3.5).

### 3.1. Exponential Discounting

We have (3.2) with  $q(l) = \int u[l(s)]c^s ds$ . The discounting parameter  $c$  can be elicited from comparison of a TTO and a SG score: Let  $a$  and  $t$  be such that  $0 < g(a) < 1$  and  $0 < h(t, a) < t$ . Then from (2.4) and the fact that  $g(a) = u(a)$ , we have

$$c^{h(t, a)} - 1 = g(a)(c^t - 1), \tag{3.5}$$

where  $c$  is the unknown.

For example if  $\widehat{h}(1, a) = 0.5$  and  $\widehat{g}(a) = 0.55$ , we obtain  $\widehat{c} \cong 0.67$ .

Thus, even a quite small difference between the SG and TTO score is an indication of high discounting. Or, the other way around, high discounting gives only rise to quite small differences in the SG and TTO scores.

In Section 3.6, plots of  $h(t, a)$  against  $u(a)$  are depicted for selected values of  $t$ . It is interesting to note that the relative distance between  $h(t, a)$  and  $u(a)t$  increases in  $t$  for all types of discounting except for power discounting where the distance is unaffected by  $t$  (see Section 3.2 below).

### 3.2. Power Discounting

We have (3.2) with  $q(l) = \int u[l(s)]s^z ds$ . Let  $a$  be a health state and  $t$  a life time where  $0 < g(a) < 1$  and  $0 < h(t, a) < t$ . Then from (2.7) and the fact that

$g(a) = u(a)$ , we have

$$h(t, a) = g(a)^{\frac{1}{z+1}} t, \quad (3.6)$$

where  $z$  is the unknown. Isolating  $z$  gives

$$z = \frac{\ln g(a)}{\ln h(1, a)} - 1.$$

For example, if  $\widehat{h}(1, a) = 0.5$  and  $\widehat{g}(a) = 0.55$  then  $\widehat{z} \cong -0.14$ .

Miyamoto and Eraker (1985) and Miyamoto (2000) show how the discounting parameter can be derived from information on ‘certainty equivalents’ which is the amount of life time in perfect health (or any other fixed health state) which is deemed equally good as some lottery with fixed probabilities of either perfect health at some positive amount of time or immediate death. The certainty equivalence (CE) method has also been used, e.g., by Gyrd-Hansen (2002) in a model with exponential discounting, by Osch *et al.* (2004) with power discounting, and by Stiggelbout *et al.* (1994) and Martin *et al.* (2000) for both exponential and power discounting.

### 3.3. Proportional Discounting

We have (3.2) with  $q(l) = \int u[l(s)] (1 + vs)^{-1} ds$ . Let  $a$  be a health state and  $t$  a life time where  $0 < g(a) < 1$  and  $0 < h(t, a) < t$ . Then from (2.9) and the fact that  $g(a) = u(a)$ , we have

$$\ln[vh(t, a) + 1] = g(a) \ln(tv + 1), \quad (3.7)$$

which may be solved numerically for  $v$ .

For example if  $\widehat{h}(1, a) = 0.5$  and  $\widehat{g}(a) = 0.55$  then  $\widehat{v} \cong 0.50$ .

### 3.4. Hyperbolic Discounting

We have (3.2) with  $q(l) = \int u[l(s)] (1 + vs)^{-w/v} ds$ , where  $v, w > 0$ . Again, the discounting parameters can be estimated as outlined in Section 2.4. We may also, slightly simpler, make use of both SG and TTO scores. Let  $a$  be a health state and  $t$  life time where  $0 < g(a) < 1$  and  $0 < h(t, a) < t$ . Then from (2.11) and the fact that  $g(a) = u(a)$  for  $v \neq w$ , we have

$$[1 + vh(t, a)]^{(-w/v+1)} - 1 = g(a)[(1 + tv)^{(-w/v+1)} - 1], \quad (3.8)$$

where  $v$  and  $w$  are unknown parameters. For a health state  $a'$ , where  $0 < g(a') < 1$  and  $g(a') \neq g(a)$ , we likewise have

$$[1 + vh(1, a')]^{(-w/v+1)} - 1 = g(a')[(1 + v)^{(-w/v+1)} - 1],$$

which is sufficient to determine  $v$  and  $w$  numerically. For completeness, it remains to consider the special case  $v = w$ . In this case, discounting is proportional (see Section 3.3).

For example, assume that  $\widehat{h}(1, a) = 0.5$ ,  $\widehat{g}(a) = 0.55$ ,  $\widehat{h}(1, a') = 0.77$ , and  $\widehat{g}(a') = 0.80$ . then

$$0.55[(1 + v)^{(-w/v+1)} - 1] = (1 + 0.5v)^{(-w/v+1)} - 1,$$

and

$$0.80[(1 + v)^{(-w/v+1)} - 1] = (1 + 0.77v)^{(-w/v+1)} - 1.$$

Solving the equations numerically yields  $\widehat{v} \cong 7.03$  and  $\widehat{w} \cong 1.56$ . The equation system is simpler than that derived from TTO scores only (Section 2.4).<sup>10</sup>

### 3.5. Constant-Proportional Risk Posture Over Exponentially Discounted Life Years

Following Johannesson *et al.* (1994), we restrict attention to chronic health states and consider constant-proportional risk posture over exponentially discounted life years. We give a separate treatment of this form of discounting to give another illustration of the implications of restricting attention to chronic health states. We have (3.2) with

$$q(l) = \left( u(a) \int_0^t c^s ds \right)^r,$$

where  $a = l(s)$ ,  $0 \leq s \leq t$ , and  $0 < c, r < 1$ . We then have

$$q(l) = [u(a)]^r \left( \int_0^t c^s ds \right)^r. \tag{3.9}$$

or

$$q(l) = \widetilde{u}(a) \left( \int_0^t c^s ds \right)^r,$$

where  $\widetilde{u}(a) = [u(a)]^r$  for all  $a$ .

What is the underlying discounting function? We are looking for a function  $\delta(s)$  satisfying

$$\int_0^t \delta(s) ds = \left( \int_0^t c^s ds \right)^r.$$

for all  $t > 0$ , where  $c$  and  $r$  are fixed parameters. As  $\left( \int_0^t c^s ds \right)^r = \left( \frac{c^t - 1}{\ln c} \right)^r$  and

$$\frac{\partial \left( \frac{c^t - 1}{\ln c} \right)^r}{\partial s} = r \left( \frac{c^t - 1}{\ln c} \right)^{r-1} c^s,$$

we obtain

$$\delta(s) = r \left( \frac{c^s - 1}{\ln c} \right)^{r-1} c^s.$$

When we restrict attention to chronic health states, this approach is therefore equivalent to the model (3.2) with a rather peculiar family of discounting functions.

We now turn to the case where health may vary over time. If we wish to preserve the property of constant-proportional risk posture (Pliskin *et al.*, 1980) over exponentially discounted life years, we have (3.2) with

$$q(t) = \int \tilde{u}[l(s)] r \left( \frac{c^s - 1}{\ln c} \right)^{r-1} c^s ds, \quad (3.10)$$

for (not necessarily chronic) health profiles  $l$ . With (3.10), the SG index  $g$  can then be used as health utility index  $\tilde{u}$ . Assume therefore that  $\tilde{u}$  is known in (3.10). We can then estimate  $r$  and  $c$  from a combination of SG and TTO scores. Let  $a$  be a health state where  $0 < g(a) < 1$  and let  $0 < h(t, a) < t$ . Then

$$g(a) \int_0^t r \left( \frac{c^t - 1}{\ln c} \right)^{r-1} c^s ds = \int_0^{h(t,a)} r \left( \frac{c^t - 1}{\ln c} \right)^{r-1} c^s ds,$$

which reduces to

$$g(a) \left( \frac{c^t - 1}{\ln c} \right)^r = \left( \frac{c^{h(t,a)} - 1}{\ln c} \right)^r, \quad (3.11)$$

or

$$r = \frac{\ln g(a)}{\ln(1 - c^{h(t,a)}) - \ln(1 - c^t)}.$$

For some health state  $a' \neq a$  with  $0 < h(t, a') < t$ , we then have

$$\frac{\ln g(a)}{\ln(1 - c^{h(t,a)}) - \ln(1 - c^t)} = \frac{\ln g(a')}{\ln(1 - c^{h(t,a')}) - \ln(1 - c^t)}.$$

This equation may then be solved numerically. For example if  $\hat{h}(1, a) = 0.5$ ,  $\hat{g}(a) = 0.55$ ,  $\hat{h}(1, a') = 0.77$ , and  $\hat{g}(a') = 0.80$  then  $\hat{c} \cong 0.88$  and  $\hat{r} \cong 0.90$ .

We have

$$h(t, a) = \frac{\ln \left\{ 1 + [g(a)]^{\frac{1}{r}} c^t - [g(a)]^{\frac{1}{r}} \right\}}{\ln c},$$

which gives

$$h'_t(t, a) = [g(a)]^{\frac{1}{r}} \frac{c^t}{1 + [g(a)]^{\frac{1}{r}} c^t - [g(a)]^{\frac{1}{r}}},$$

which is positive, and after some calculations, we obtain

$$h''_{tt}(t, a) = (\ln c) c^t \frac{[g(a)]^{\frac{1}{r}} - [g(a)]^{\frac{2}{r}}}{\left\{ 1 + [g(a)]^{\frac{1}{r}} c^t - [g(a)]^{\frac{1}{r}} \right\}^2},$$

which is negative. We therefore find that  $h(t, a)$  is also strictly concave in  $t$ .

A different approach is to consider a model (3.2) with

$$q(l) = \left\{ \int \tilde{u}[l(s)]e^s ds \right\}^r, \tag{3.12}$$

for (not necessarily chronic) health profiles  $l$ . In this case, we give up constant-proportional risk posture over exponentially discounted life years and obtain a model that is not of the form (3.3) but captures a (power) form of risk aversion over exponentially discounted QALYs. We provide some further discussion of this type of models in Section 4.

3.6. Identification of an Appropriate Discounting Model

As is the case for deterministic QALY models, it will likewise be possible to extract information on the appropriate discounting model by designing empirical investigations. The previous sections show that mathematical relationships exist between TTO and SG scores and the discounting parameter(s) of the five discounting models [equations (3.1), (3.2), (3.3), (3.8), and (3.11)]. Figures 5–7 display plots of pairs of  $[g(a), h(t,a)]$  for fixed  $t$  and specific discounting parameter(s). We use the same parameters as in Figures 1–4. We consider  $t = 1$  (Figure 5),  $t = 10$  (Figure 6) and  $t = 50$  (Figure 7). Empirical values of  $g(a)$  and  $h(t, a)$  must be collected for health states sufficiently different so that the resulting set of values  $g(a)$  spans most of the interval  $[0, 1]$ . For each health state  $a$  included in

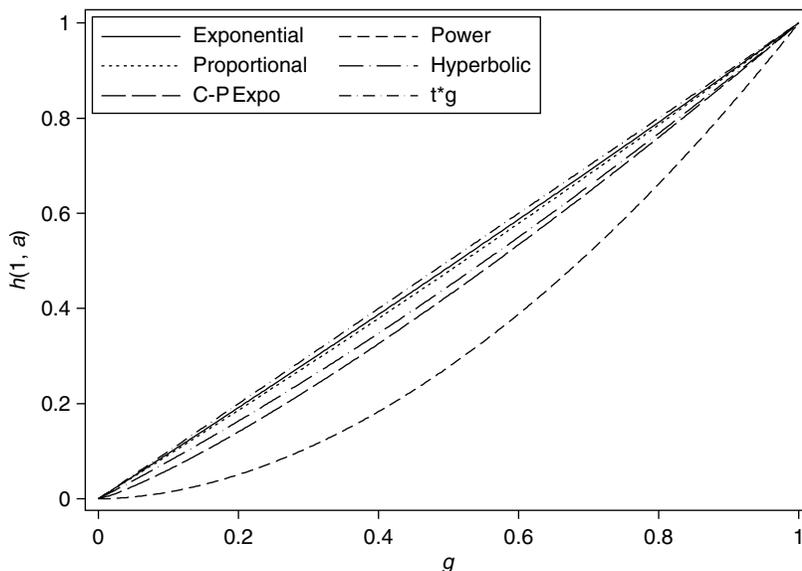


Figure 5. Plot of  $h(1, a)$  against  $g(a)$ .

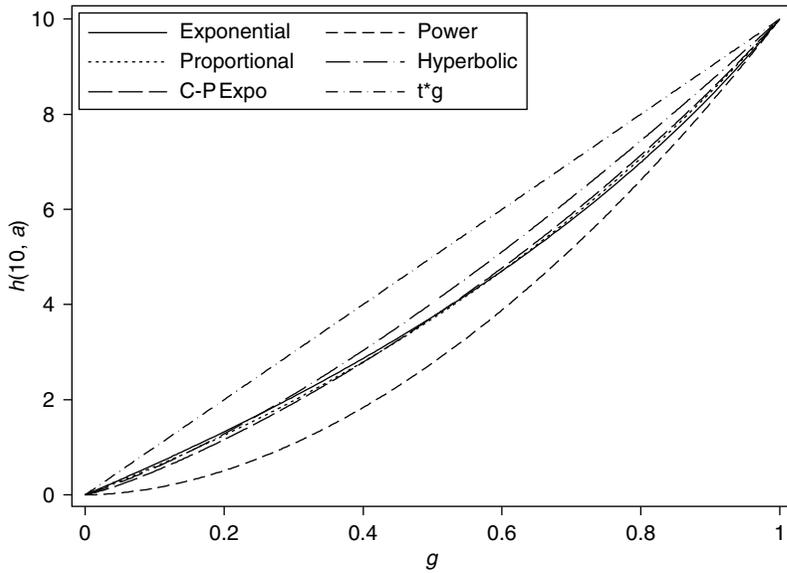


Figure 6. Plot of  $h(10, a)$  against  $g(a)$ .

the set, there will then be a pair of values  $[g(a), h(t, a)]$ . If these pairs of health state values for an individual are plotted into a  $[g(a), h(t, a)]$  diagram, it may be possible to deduce a possible discounting model. Most probably, the utilization of

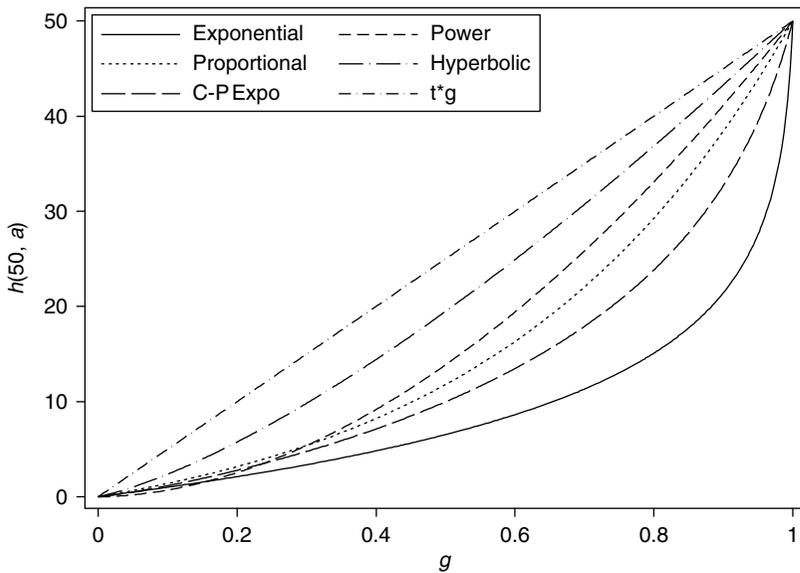


Figure 7. Plot of  $h(50, a)$  against  $g(a)$ .

econometric methods to identify the best discounting model and parameter(s) will be needed.

It is worth noticing the implication of the (generalized) constant proportional trade-off property of power discounting. Under power discounting, the relative distance between the curves  $h(t, a)$  and  $tu(a)$  is independent of  $t$ . This is in stark contrast to, for instance, exponential discounting, where the curvature of  $h(t, a)$  varies strongly with the time horizon.

#### 4. Discussion

For the following discussion, we recall that *risk aversion* relates to the curvature of the function  $f$  when we have a representation of the form (3.4), whereas *discounting* relates to the functional form of  $\delta$  in a model which can either be of the type (3.3) or (3.4).

##### 4.1. Is SG the gold standard?

In the literature, some authors have claimed that QALYs are utilities while others have disputed this. In this paper, the QALY model  $q$  (or  $Q$ ) represents a preference relation on the space of (lotteries over) health profiles, and as such, the QALY model is by construction a utility function in the health domain.

The SG method has often been called the ‘gold standard.’ Drummond *et al.* (1997), for example, write in their book that ‘A utility, in our context here, is a von Neumann-Morgenstern utility. So all QALYs that are formed from preferences measured in any other way other than with a SG, by definition, can not be utilities’ (p. 183). The authors do not refer to any explicitly stated model, and it is interesting to revisit this conclusion using the framework of Section 3.

First, suppose that the utility representation is of the form (3.3), a model with risk neutrality over QALYs. In this case, the utility function may be estimated fully with TTO scores only. Once the discounting function has been determined (from TTO scores with varying time horizon), the health state index is not equal to the TTO score but uniquely determined by it which is sufficient for our purpose. In fact, contrary to the TTO score, the SG scores cannot be used for eliciting the discounting function, and in this sense, they are less useful than TTO scores.<sup>11</sup>

Second, suppose on the contrary that risk neutrality over QALYs is not necessarily assumed and the underlying model is (3.4) for some strictly increasing concave function  $f$ . (Perhaps this more general model fits better with the discussion in Section 6.2.2 of Drummond *et al.*). Typically, it would then be assumed that  $f$  is a member of some parametric family with one of two free parameters that can be estimated with SG scores. However, the shape of the function  $v$  is not related to preferences over lotteries and possible risk aversion, and  $v$  is naturally estimated from TTO scores.

To conclude, regardless of the specific QALY model, TTO scores are indeed relevant for parameter estimation also in an expected utility framework.

#### 4.2. *The Role of the Constant Proportional Trade-Off Assumption*

In a paper discussing the SG method, Gafni (1994, p. 211) argues that the SG method requires constant proportional trade-off.<sup>12</sup> However, at a closer inspection, this claim cannot be correct. In any model of the form (3.3) with  $v(a, s) = u(a)\delta(s)$ , where constant proportional trade-off is not necessarily satisfied, the SG scores can be substituted directly for the health state index.

Generally, it is useful to clearly distinguish between the underlying model and the particular method used for parameter estimation. The SG method can be useful for parameter estimation in *any* model assuming of course that we can rely on the procedure used for estimating SG scores (an empirical question which is likely to depend on the context). It may not be the case that the SG estimate can be used directly as a health state index as it for example is the case with risk aversion over QALYs. This is similar to TTO scores which must be transformed depending on the specific family of discounting models used but does not make the scores less applicable for parameter estimation.

Broome (1993) argued in a critique of the axioms introduced by Pliskin *et al.* (1980) that constant proportional trade-off rules out any discounting of future QALYs and is out of place at the level of a general theory. We shall not argue about the appropriate level of generality in health economic theory, but we may notice that the constant proportional trade-off is consistent with the family of power discounting functions, a property that might be useful in applications. Indeed, constant-proportional trade-off is a powerful assumption with strong implications. On the contrary, it is rather easily understood and is simple to 'test' in stated preference experiments. The family of power discounting functions may be sufficiently rich for some purposes, but this clearly depends on the nature and quality of available data.

#### 4.3. *The Problem of 'Double Discounting of QALYs'*

We have seen in Sections 2 and 3 that under discounting of any form, the ratio between the TTO score and the time horizon is different from the health state index, and the TTO score must be adjusted for in a way that depends on the type and degree of discounting. If an exogenous discount factor is used, but the TTO score/time horizon ratio nevertheless is used as a health state index, the health state values are underestimated for TTO scores strictly between zero and the time horizon. The absolute number of QALYs is accordingly also smaller relative to the situation where the same discount factor and the correct health index is used. This phenomenon is referred to as 'double discounting of QALYs.'

Krahn and Gafni (1993) argue that discounting QALYs may result in double discounting both in case of the SG and the TTO method, because time preference is already incorporated in utility assessment (p. 413–414). As the SG score is not affected by discounting in models of the form (3.3) with  $v(s, a) = u(a)\delta(s)$  (which are standard in the context of QALYs), the 'double discounting' referred to by the authors must be a bias of different nature than the above mentioned in the case of the SG method. For instance, in choice experiments, respondents may not act in

accordance with expected utility. ‘Double discounting of QALYs’ is potentially a misleading terminology in this case, because if the SG estimates would be biased, it would be for reasons that have nothing to do with the problem of discounting twice.

The issue of double discounting is not the use of a discount factor in conjunction with TTO-based health indices but the failure of adjusting the TTO score/time horizon ratio before using it as health index. The real problem is that failing to adjust would tend to underestimate bad health states and overestimate good health states, because the *relative* distance between  $h(t, a)$  and  $u(a)t$  decreases in  $u(a)$  as illustrated in Figures 5–7 in Section 3.6. Consequently, there seems to be little theoretical justification for attempting to adjust for double discounting by reducing or entirely eliminating discounting which seems to be the idea in a paper by MacKeigan *et al.* (2003).<sup>13</sup>

#### 4.4. *TTO, SG, or Both?*

There is a large and growing literature examining the relative advantages of TTO and SG methods (or other probabilistic methods) usually taking the point of departure in data gathered from questionnaire studies. For selected references, see Bleichrodt (2002) and Dolan (2000).

In a recent paper, for example, Gyrd-Hansen (2002) reports a study where time preferences were elicited through TTO scores and also with CE scores. It was predicted and confirmed experimentally that the discount rate elicited through TTO scores would be lower than those elicited through the CE scores, as the latter incorporate risk attitude (as with the SG score) as opposed to the TTO method. From this it was concluded that the TTO method is the most troublesome (compared with the CE method).

It seems plausible that the discrepancy between discount factors estimated from these two methods is explained by the fact that the CE method is affected by risk aversion with respect to discounted quality-adjusted life years, but we propose a different (or additional) conclusion from this observation. We could (should) use *both* TTO and SG estimates (or certainty equivalence methods for that matter) to identify risk aversion and disentangle this from discounting. If the model is sufficiently flexible and allows for risk aversion, the TTO scores are not less useful than SG scores.

Broome (1993) criticizes the SG method for not providing the right answer (to the health state index) in case of risk aversion over QALYs. In the example he provides, health state preferences are of the form (3.12), with  $r = 0.5$  and with no discounting of deterministic health profiles. Johannesson (1995) argued in a comment to Broome’s paper that the health state index and the parameter  $r$  can be derived from a combination of SG and TTO scores and outlined a procedure similar to that in Section 3.2. However, Johannesson’s argument relies on the assumption of chronic health states using the fact that in this case the model can be restated in the form (3.9).

With non-chronic health, there is an important difference between risk aversion and discounting, and typically, neither the risk parameter  $r$  nor the discounting parameter(s) are known from the outset and must be estimated from observations as with the health state index. Again, we need both TTO and SG estimates: TTO scores can be used for estimating the discounting function and the health state index and SG estimates can be used to identify the risk parameter.

When all health states are chronic, risk aversion and discounting may be exactly the same thing as illustrated in Section 3.5. But it is crucial to notice that this is only a special feature of some models with chronic health states that does not carry over to models where health is allowed to vary over time.

## 5. Conclusion and Final Remarks

The paper has provided an overview of QALY models and derived procedures for parameter estimation from the type of data that is often collected in QALY studies. The models deal with varying health states, but with one exception (discussed in Section 2.2), procedures involving only comparisons of certain health profiles with chronic health states can be devised. A preliminary discussion of model identification was also provided, although our treatment of this important issue was very incomplete.

The appendix provides an overview of examples of previous studies which have used experimental data to estimate parameter(s) in one or more of the specific QALY models presented in the previous sections.

Of course, the choice of model and the number of free parameters should reflect the availability and quality of data. In any case, however, time preference is an integral part of preference for health profiles, and accordingly, the estimation of discount factors is an integral part of health preference measurement. We hope that the present survey has contributed to a further elaboration of the implications of this point for the use of TTO- and SG-based procedures.

Finally, we mention three important limitations of our analysis. First, many experiments have indicated that respondents' judgements of probability are not linear in probability (see, e.g., Kahneman and Tversky, 1979; Camerer, 1995; Gonzalez and Wu, 1999). In this paper, we have restricted attention to deterministic models and expected utility models. Wakker and Stiggelbout (1995) and more recent papers by Bleichrodt and Pinto (2000), and Bleichrodt *et al.* (2001) have developed procedures that might be combined with methods outlined in Section 3, but an investigation of this in a QALY context is beyond the scope of the present paper (see also Verhoef *et al.*, 1994; Miyamoto, 2000; Bleichrodt, 2002; Osch *et al.*, 2004; Bleichrodt and Pinto, 2005).

Second, in many applications, TTO and SG scores are obtained from a group of respondents and the scores are aggregated in some way to form the preferences of one 'representative' individual. We have not addressed questions related to interpersonal aggregation, but of course the TTO and SG scores assumed to be available could each represent a mean or a median of a sample of scores.

Third, we have already mentioned the need for developing statistical techniques for the analysis of 'noisy' data. For this, a random utility extension of the present framework would seem appropriate. The micro-econometrics of QALYs in a framework as the one presented in this paper appears to be largely unexplored, and we suggest that development of random utility QALY models and from this derivation of relevant statistical methods for model identification and parameter estimation in this specific context would be an interesting area for future research.

### Acknowledgements

This paper is a revised version of Discussion Paper 04-05 'QALYs When Health Varies Over Time: An Analysis of Model Identification and Parameter Estimation from Time Trade-Off and Standard Gamble Scores', Institute of Economics, University of Copenhagen, 2004. The authors are grateful to an anonymous referee and Editor Les Oxley for very helpful comments.

### Notes

1. The reasonableness of this assumption depends on the interpretation of the health states. For choice experiments and discussions, see the articles referred to in Section 1.
2. The integrals we consider in this paper are well defined under mild technical assumptions, see Grodal (2003).
3. For a characterization of this class of functions in terms of conditions on the preference relation, see Grodal (2003, Section 13.3).
4. We have used the parameters  $c = 0.9$  (for the exponential discounting model),  $z = -0.4576$  (power discounting),  $v = 0.1868$  (proportional discounting), and  $v = 4$ ,  $w = 1.1349$  (hyperbolic discounting).
5. We have  $\int (1 + vs)^{-w/v} ds = \frac{(1+vs)^{(-w/v+1)}}{(-w/v+1)v}$ ,  $v \neq w$ .
6. It is beyond the scope of this paper to investigate efficient ways of solving equation systems of this type. Here, we confine ourself to note that using numerical solution methods it seems that non-valid solutions (for which  $\hat{v} = \hat{w}$ ) are likely to appear unless the initial guess value is fairly close to the valid solution.
7. For example, if  $u(a) = 0.5$ ,  $v = 4$ , and  $w = 1.1349$ , we have  $h''_{tt}(1, a) \cong -0.023$ ,  $h''_{tt}(10, a) \cong -0.00071$  and  $h''_{tt}(100, a) \cong -0.000015$ .
8. For example, if  $u(a) = 0.5$ ,  $v = 1$ , and  $w = 1.1$  then  $h(t, a)$  converges toward 1023.0, but we have  $h(10, a) \cong 2.0871$ ,  $h(1000, a) \cong 16.624$ ,  $h(10^{10}, a) \cong 393.8$  and  $h(10^{30}, a) \cong 1012.8$  etc.
9. The parameters used are  $c = 0.95$  and  $r = 0.8286$  (such that the discount factor is equal to the other discount factors at  $s = 10$ ).
10. Note that solutions with  $\hat{v} = \hat{w}$  are not valid in this equation system (see also note 5).
11. However, the discounting factor can be derived from 'certainty equivalents' described in Section 3.2.
12. This condition means that for each health state  $a$  there exists a number  $\lambda$  such that  $t$  years in  $a$  is equivalent to  $\lambda t$  years in perfect health  $a^*$  for all  $t > 0$ .
13. In the study by MacKeigan *et al.*, the authors aim to determine the magnitude of the double discounting effect by comparing an undiscounted holistic TTO score (obtained from the evaluation of a non-chronic health profile) with composite scores (obtained

from separate evaluations of chronic health states) explicitly discounted by discounting rates 0%, 3%, or 5%. The 0% discounted composite score was closest to the undiscounted holistic score, but this is not necessarily evidence of double discounting. (For example think of the situation where health states are relatively similar and discounting is low.)

## References

- Bleichrodt, H. (2002). A new explanation for the difference between time trade-off utilities and standard gamble utilities. *Health Economics* 11: 447–456.
- Bleichrodt, H. and Johannesson, M. (1997). Standard gamble, time trade-off and rating scale: experimental results in the ranking properties of QALYs. *Journal of Health Economics* 16: 155–175.
- Bleichrodt, H. and Johannesson, M. (2001). Time preference for health: a test of stationarity versus decreasing timing aversion. *Journal of Mathematical Psychology* 45: 265–282.
- Bleichrodt, H. and Pinto, J. L. (2000). A parameter-free elicitation of the probability weighting function in medical decision analysis. *Management Science* 46: 1485–1596.
- Bleichrodt, H. and Pinto, J. L. (2005). The validity of QALYs under non-expected utility. *Economic Journal* 115: 533–550.
- Bleichrodt, H., Pinto, J. L. and Wakker, P. (2001). Making descriptive use of prospect theory to improve the prescriptive use of expected utility. *Management Science* 47: 1498–1514.
- Bleichrodt, H., Wakker, P. and Johannesson, M. (1997). Characterizing QALYs by risk neutrality. *Journal of Risk and Uncertainty* 15: 107–114.
- Broome, J. (1993). Qalys. *Journal of Public Economics* 50: 149–167.
- Cairns, J. A. and van der Pol, M. M. (1997a). Constant and decreasing timing aversion for saving lives. *Social Science and Medicine* 45: 1653–1659.
- Cairns, J. A. and van der Pol, M. M. (1997b). Saving future lives: a comparison of three discounting models. *Health Economics* 6: 341–340.
- Cairns, J. A. and van der Pol, M. M. (2000). Valuing future private and social benefits: the discounted utility model versus hyperbolic discounting models. *Journal of Economic Psychology* 21: 191–205.
- Camerer, C. (1995). Individual decision making. In J. H. Kagel and A. E. Roth (eds), *The Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press.
- Cher, D., Miyamoto, J. and Lenert, L. (1997). Incorporating risk attitude into Markov-process decision models: importance for individual decision making. *Medical Decision Making* 17: 340–350.
- Doctor, J. and Miyamoto, J. (2003). Deriving quality-adjusted life years (QALYs) from constant proportional time tradeoff and risk posture conditions. *Journal of Mathematical Psychology* 47: 557–567.
- Dolan, P. (2000). The measurement of health-related quality of life. In A. Culyer and J. P. Newhouse (eds), *Handbook of Health Economics*. Amsterdam: Elsevier, chapter 32.
- Dolan P. and Gudex C. (1995). Time preference, duration and health state valuations. *Health Economics* 4: 289–299.
- Dolan, P., Gudex, C., Kind, P. and Williams, A. (1996). Valuing health states: a comparison of methods. *Journal of Health Economics* 15: 209–231.
- Drummond, M. F., O'Brien, B. J., Stoddart, G. L. and Torrance, G. W. (1997). *Methods for the Economic Evaluation of Health Care Programmes*, 2nd edn. Oxford University Press.
- Gafni, A. (1994). The standard gamble method: What is being measured and how it is interpreted. *Health Services Research* 29: 207–224.

- Gafni, A. and Torrance, G. (1984). Risk attitude and time preference in health. *Management Science* 30: 440–451.
- Gonzalez, R. and Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology* 38: 129–166.
- Grodal, B. (2003). Integral representations. In K. Vind (ed.), *Independence, Additivity, Uncertainty*. Berlin: Springer, chapters 12–13.
- Gyrd-Hansen, D. (2002). Comparing the results of applying different methods of eliciting time preferences for health. *European Journal of Health Economics* 3: 10–16.
- Harvey, C. (1986). Value functions for infinite-period planning. *Management Science* 32: 1123–1139.
- Harvey, C. (1995). Proportional discounting of future costs and benefits. *Mathematics of Operations Research* 20: 381–399.
- Herrnstein, R. J. (1981). Self-control as response strength. In C. M. Bradshaw, E. Szabadi and C. F. Lowe (eds), *Quantification of Steady-State Operant Behavior*. Amsterdam: Elsevier, North-Holland.
- Johannesson, M. (1995). Qalys: a comment. *Journal of Public Economics* 56: 327–328.
- Johannesson, M., Pliskin, J. and Weinstein, M. (1994). A note on QALYs, time tradeoff, and discounting. *Medical Decision Making* 4: 188–193.
- Kahneman, D. and Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica* 47: 263–291.
- Koopmans, T. C. (1972). Representation of preference orderings over time. In C. B. McGuire and R. Radner (eds), *Decision and Organization*. Amsterdam: North-Holland.
- Krabbe, P. F. M. and Bonsel, G. J. (1998). Sequence effects, health profiles, and the QALY model: in search of a realistic model. *Medical Decision Making* 18: 178–186.
- Krahn, M. and Gafni, A. (1993). Discounting in the economic evaluation of health care interventions. *Medical Care* 31: 403–418.
- Kreps, D. (1988). *Notes on the Theory of Choice*. Boulder: Westview Press.
- Kuppermann, M., Shiboski, S., Feeny, D., Elkin, E. P. and Washington, A. E. (1997). Can preference scores for discrete states be used to derive preference scores for an entire path of events? *Medical Decision Making* 17: 42–55.
- Lipscomb, J. (1989). Time preference for health in cost-effectiveness analysis. *Medical Care* 27: 233–253.
- Loewenstein, G. and Prelec, D. (1992). Anomalies in intertemporal choice: evidence and an interpretation. *Quarterly Journal of Economics* 107: 573–597.
- MacKeigan, L., Gafni, A. and O'Brien, B. (2003). Double discounting of QALYs. *Health Economics* 12, 165–169.
- Martin, A. J., Glasziou, P. P., Simes, R. J. and Lumley, T. (2000). A comparison of standard gamble, time trade-off, and adjusted time trade-off scores. *International Journal of Technology Assessment in Health Care* 16: 137–147.
- Mazur, J. E. (1987). An adjustment procedure for studying delayed reinforcement. In M. L. Commons, J. E. Mazur, J. A. Nevins and H. Rachlin (eds), *Quantitative Analysis of Behavior: The effect of Delay and of Intervening Events on Reinforcement Value*. Hillsdale: Ballinger, chapter 2.
- Miyamoto, J. (1999). Quality-adjusted life years (QALY) utility models under expected utility and rank-dependent utility assumptions. *Journal of Mathematical Psychology* 43: 201–237.
- Miyamoto, J. (2000). Utility assessment under expected utility and rank-dependent utility assumptions. In G. Chapman and F. Sonnenberg (eds), *Decision Making in Health Care: Theory, Psychology, and Applications*. Cambridge University Press.
- Miyamoto, J. and Eraker, S. (1985). Parameter estimates for a QALY utility model. *Medical Decision Making* 5: 191–213.

- Miyamoto, J. and Eraker, S. (1988). A multiplicative model of the utility of survival duration and health quality. *Journal of Experimental Psychology: General* 117: 3–20.
- Miyamoto, J. and Eraker, S. (1989). Parametric models of the utility of survival duration. *Organizational Behaviour and Human Decision Processes* 44: 166–202.
- Miyamoto, J., Wakker, P., Bleichrodt, H. and Peters, H. (1998). The zero-condition: a simplifying assumption in QALY measurement and multiattribute utility. *Management Science* 44: 839–849.
- Olsen, J. A. (1994). Person vs years: two ways of eliciting implicit weights. *Health Economics* 3: 39–46.
- Osch, S., Wakker, P., van den Hout, W. and Stiggelbout, A. (2004). Correcting biases in standard gamble and time tradeoff utilities. *Medical Decision Making* 24: 511–517.
- Pliskin J., Shepard, D. and Weinstein, M. (1980). Utility functions for life years and health status. *Operations Research* 28: 206–224.
- van der Pol, M. M. and Cairns, J. A. (2002). A comparison of the discounted utility model and the hyperbolic discounting model in the case of social and private intertemporal preferences for health. *Journal of Economic Behavior & Organization*. 49: 79–96.
- van der Pol, M. M. and Roux, L. (2005). Time preference bias in time trade-off. *European Journal of Health Economics* 6: 107–111.
- Redelmeier D. A. and Heller D. N. (1993). Time preference in medical decision making and cost-effectiveness analysis. *Medical Decision Making* 13: 212–217.
- Richardson, J., Hall, J. and Salkeld, G. (1996). The measurement of utility in multiphase health states. *International Journal of Technology Assessment in Health Care* 12: 151–162.
- Spencer, A. (2003). A test of the QALY model when health varies over time. *Social Science and Medicine* 57: 1697–1706.
- Stavem, K., Kristiansen I. S. and Olsen, J. A. (2002). Association of time preference for health with age and disease severity. *European Journal of Health Economics* 3: 120–124.
- Stiggelbout, A. M., Keibert, G. M., Keivit, J., Leer, J. W. H., Stoter, G. and de Haes, J. C. J. M. (1994). Utility assessment in cancer patients: adjustment of time tradeoff scores for the utility of life years and comparison with standard gamble scores. *Medical Decision Making* 14: 82–90.
- Torrance, G. W. (1986). Measurement of health state utilities for economic appraisal. *Journal of Health Economics* 5: 1–30.
- Treadwell, J. (1998). Tests of preferential independence in the QALY model. *Medical Decision Making* 18: 418–428.
- Verhoef, L., de Haan, A. and van Daal, W. (1994). Risk attitude in gambles with years of life: empirical support for prospect theory. *Medical Decision Making* 15: 180–186.
- Wakker, P. and Stiggelbout, A. (1995). Explaining distortions in utility elicitation through the rank-dependent model for risky choices. *Medical Decision Making* 15: 180–186.
- Zweifel, P. and Breyer, F. (1997). *Health Economics*. Oxford University Press.
- Østerdal, L. P. (2005). Axioms for health care resource allocation. *Journal of Health Economics* 24: 679–702.

## Appendix

### Literature with Joint Estimation of Discounting Parameter and Health State Index

| Elicitation method   | Discounting model(s) applied | Discounting parameter estimated (mean value)               | Source                           |
|--|------------------------------|--|----------------------------------|
| CE for discounting parameter, TTO for health state index       | Power                        | Power coefficient 1.14                                     | Pliskin <i>et al.</i> (1980)     |
| CE for discounting parameter, TTO for health state index       | Power                        | Power coefficient: 1.03                                    | Miyamoto and Eraker (1985)       |
| TTO or RS for discounting parameter and health state index     | Exponential                  | Discount rate: 2.3–3.9% (depending on health state)        | Redelmeier and Heller (1993)     |
| TTO for discounting parameter and health state index           | Exponential                  | Discount rate: 10.0%                                       | Olsen (1994)                     |
| CE for discounting parameter, SG or TTO for health state index | Power                        | Power coefficient: 0.74                                    | Stiggelbout <i>et al.</i> (1994) |
| TTO for discounting parameter and health state index           | Exponential                  | Discount rate –3.5%–1.4% (depending on health state)       | Dolan and Gudex (1995)           |
| CE for discounting parameter, SG or TTO for health state index | Exponential                  | Discount rate: 9.0%–18.0% (depending on survival duration) | Martin <i>et al.</i> (2000)      |
| TTO with non-chronic health states for discounting parameter   | Exponential                  | Discount rate: 0.9%–2.4% (depending on duration)           | van der Pol and Roux (2005)      |
| TTO for health state index                                     | Power                        | 'Time preference rate': 7.5%–20.0% (depending on duration) |                                  |

CE: certainly equivalence; RS: rating scale; SG: standard gamble; TTO: time trade-off.

Notes: This appendix provides a table with examples of previous literature which has used experimental data for the joint estimation of a discounting parameter and a health state index in specific QALY models.

