In G. B. Chapman and F. Sonnenberg, *Decision making in health care: Theory, psychology, and applications*. New York: Cambridge University Press, 65-109.

Utility Assessment under Expected Utility and Rank Dependent Utility Assumptions

John M. Miyamoto
University of Washington

Printed Tuesday, March 19, 2002

Acknowledgments

I would like to thank Richard Gonzalez, Jonathan Treadwell, and Peter Wakker for helpful discussions of issues in utility theory and utility assessment, and Han Bleichrodt for commenting on an earlier version of this manuscript.

Abstract

Although the methodology of utility assessment under expected utility assumptions is well understood, empirical research has demonstrated that expected utility theory is not descriptively valid. Accordingly, new methods of utility assessment must be found that are consistent with the human preference behavior. This chapter presents methods of utility assessment under assumptions of rank dependent utility theory. Rank dependent utility theory models the nonlinear perception of probabilities, which is a major source of violations of expected utility theory. It is shown that the standard gamble method, the time tradeoff method, and the method of certainty equivalents can all be interpreted from the standpoint of rank dependent utility theory. Furthermore, it is shown that utility assessments under rank dependent utility assumptions differ systematically from utility assessments under expected utility assumptions. In particular, measures of risk aversion and of health state utilities are systematically affected by the nonlinear perception of probability.

Utilities have become a standard measure of value in the analysis of health decisions (Drummond, O'Brien, Stoddart, & Torrance, 1997; Gold, Siegel, Russell, & Weinstein, 1996). For purposes of decision analysis, utilities have a number of highly desirable properties: They are grounded in a normative theory of preference whose mathematical and theoretical foundations are well understood (Fishburn, 1982, 1989; Edwards, 1992; von Neumann & Morgenstern, 1944); they allow the construction of preference models that are adapted to the structure of specific decisions and outcome domains (Keeney & Raiffa, 1976; Sox, Blatt, Higgins, & Marton, 1988; von Winterfeldt, & Edwards, 1986); and they possess a well developed methodology for measuring the utilities of health outcomes as perceived by relevant populations of individuals (Keeney & Raiffa, 1976; Weinstein et al., 1980). It would seem that the stage is set for the unfettered application of utility theory to the task of assessing the value of health. There is, however, a serious obstacle to proceeding down this path. Extensive empirical research has demonstrated a variety of ways in which preferences are inconsistent with the assumptions of expected utility theory (Kahneman & Tversky, 1979, 1984; Slovic, Lichtenstein, & Fischhoff, 1988). If the methodology of utility assessment is largely based on expected utility theory, a theory that can be rejected as a descriptive theory of preference, how are utilities to be measured?

This is a big question, and its answer is still very much under development. What I hope to do here is present one line of attack on the assessment of utilities that is based on rank dependent utility theory. Rank dependent utility theory postulates that the probabilities as stated in lotteries do not directly determine the utility of lotteries. Rather the probabilities are first transformed nonlinearly to decision weights, which are then combined with outcome utilities to determine the utility of lotteries (Quiggin, 1982; Karni & Safra. 1990). The nonlinear transformation of probabilities introduces fundamental changes in the methodology of utility assessment (Wakker & Stiggelbout, 1995). This chapter presents the Wakker/Stiggelbout analysis of the standard gamble method under rank dependent utility assumptions, and extends their analysis to problems in QALY measurement and the characterization of risk posture. The chapter also describes the implications of nonlinear probability perception for the interpretation of certainty equivalents and time tradeoffs. Many utility assessment procedures that are standard in medical decision making assume the validity of expected utility theory, and hence, they assume that the perception of probability is linear. What I hope to do is to describe how the nonlinear perception of probability distorts utilities that are assessed by standard methods, and to describe methods for removing the effects of these distortions.

Before undertaking this discussion, I should try to be clear about the perspective on utility analysis taken in this chapter. The focus of this chapter is on the theory that underlies *utility assessment procedures*, in other words, procedures by which numbers representing values are assigned to health outcomes. Of course, I do not have in mind just any procedures, but rather those that can be justified from the standpoint of a theory of preference under risk. I assume the normative validity of expected utility theory (henceforth, EU theory), and assume that the goal of utility assessment is to provide a quantitative measure of preference that can be combined with probabilities in a utility analysis of health decisions or policies (Gold et al., 1996; Weinstein et al., 1980). Because the descriptive validity of EU theory is no longer tenable, a need has arisen for utility assessment procedures that take violations of EU theory into account.

There are four major findings in the psychology of preference that must be taken into account in the theory of utility assessment. The first has already been mentioned, namely, the nonlinear perception of probability. Second, it has been argued that people represent outcomes as gains or losses relative to a neutral reference level, rather than as absolute states of wealth (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). The categorization of outcomes as gains or losses has predictable effects on risk posture and the rate at which utility changes as a function of objective changes (so-called loss aversion). Third, preferences are affected by the way in which choices are framed (Kahneman & Tversky, 1984). Finally, preferences as inferred from choices are not identical to preferences as inferred from matching tasks like judgments of selling prices or certainty equivalents (Slovic, & Lichtenstein, 1968; Tversky, Sattath, & Slovic, 1988; Fischer, G. W., & Hawkins, S. A., 1983; Bostic, Herrnstein, & Luce, 1990). This essay is primarily an attempt to incorporate nonlinear probability perception into the methodology of utility assessment. The other issues will not be addressed in this essay.

For simplicity, this chapter will focus on the problem of assessing the utility of health states that are better than death. A better-than-death health state is a state in which longer survival is preferred to shorter survival. Not all health states are better than death. Some health states are worse than death (shorter survival in these states is preferred to longer survival), and some health states are regarded as having a maximum endurable survival (longer survival is preferred to shorter survival up to the point of maximum endurable survival, and then shorter survival is preferred to longer survival beyond this point). Such health states give rise to important and interesting assessment issues, but these issues would digress from the central questions of this chapter. Patrick, Starks, Cain, Uhlmann, and Pearlman (1994) and Drummond et al. (1996) discuss the problem of assessing worse than death health states. Sutherland, H. J., Llewelynn-Thomas, Boyd, and Till (1982) pointed out the occurrence of maximum endurable survivals, and Stalmeier, Bezembinder, and Unic (1996) noticed some problematic inconsistencies in judgments involving maximum endurable survivals.

This chapter has the following organization. The first section of the paper describes four basic problems in utility assessment. These problems are the assessment of holistic outcomes, the assessment of a utility function for survival duration, the assessment of a linear QALY model, and the assessment of a power QALY model. The second section reviews how EU theory solves these assessment problems by means of standard gambles, certainty equivalents, and time tradeoffs. Although this material is well known, the discussion will emphasize those features of assessment procedures that must be revised when nonlinear probability weighting is taken into account. The third section presents solutions to the utility assessment problems under the assumptions of rank dependent utility theory. The goal of this section is to show how standard utility assessments must be reinterpreted in a framework that allows for the nonlinear weighting of probability. The fourth section presents some preference data, and compares an analysis from the EU standpoint to an analysis from the rank dependent utility standpoint. The section shows concretely how nonlinear probability weighting affects the interpretation of health preference data. The results presented in the third and fourth sections are the main contribution of this chapter. A final section reviews the problem of utility assessment from a perspective that takes violations of EU theory into account.

Four Problems in Utility Assessment

Before discussing the theory of utility assessment, we should list the types of assessment problems that this theory is intended to solve. As I describe these assessment problems, I will assume that we are trying to determine the utilities of a specific person, who I will refer to as "the client." Depending on the research problem, the client may be a patient, a health professional, or a person drawn from the general public.

The conceptually simplest problem is one in which one has a short list of distinct health outcomes whose utility is to be assessed. For example, Sox et al. (1988) discuss a decision between surgical and medical treatment of a herniated intervertebral disc that is causing severe back pain. The potential outcomes in this decision are complete recovery, residual back pain, perioperative death, and residual back pain with footdrop. In this case, the utility assessment problem is to determine utilities that represent the relative worth for the client of each of these outcomes.

Utility Analysis (UA) Problem 1. Given a finite list of health outcomes, *A*, *B*, *C*, ..., determine utilities, *U(A)*, *U(B)*, *U(C)*, ..., for the outcomes on this list.

This type of utility assessment problem is sometimes called an assessment of *holistic* outcomes because the analysis does not attempt to decompose the outcomes into attributes whose separate utilities are assessed and then combined by a composition rule.

Consider, next, the case where duration of survival is a component of the possible health outcomes. In this case, one usually attempts to represent the utility of survival duration by means of a smooth curve that represents the increase in utility as a function of duration.

UA Problem 2. Given that health outcomes lie on a continuum like survival duration, construct a curve that approximates the growth in utility along this continuum for a particular client.

Before proceeding to other assessment problems, it is worthwhile to digress briefly on the issue of risk posture. *Risk posture* or *risk attitude* refers to the curvature of a utility function over a continuum like survival duration or money. Technically, when talking about risk posture, one must

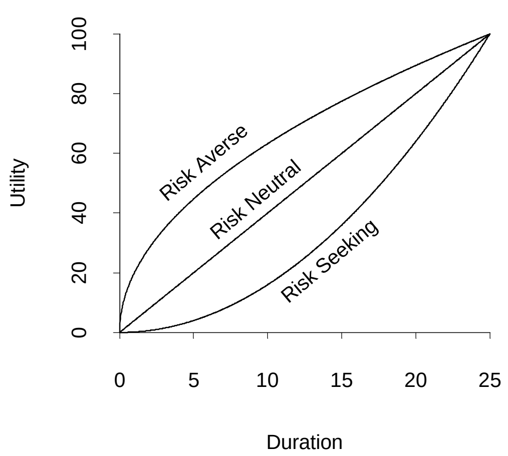

Figure 1. Risk averse, risk neutral, and risk seeking utilities functions for survival duration.

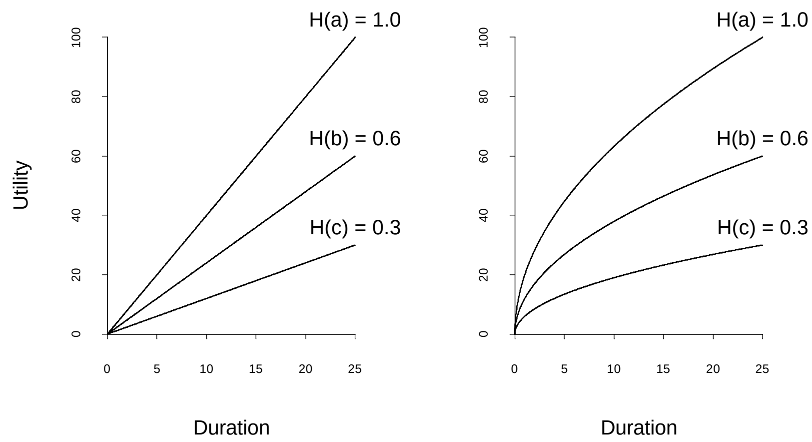

Figure 2. Left panel: The (linear) QALY model with three health states. Right panel: Power QALY model with three health states (power parameter r = .5).

specify the continuum in question, for it is perfectly reasonable (logically consistent) to have different risk postures with respect to survival duration, money, number of lives saved, etc. Figure 1 displays utility functions for survival duration that exhibit different risk postures with respect to survival duration. A utility function is *risk averse* if it is concave downward over the continuum in question. It is *risk seeking* if it is concave upwards over the continuum, and it is *risk neutral* if it is linear over the continuum. Risk posture is important in medical decision making because therapies can differ in their tradeoffs between short and long term survival (McNeil, Weichselbaum, & Pauker, 1978; McNeil & Pauker, 1982; Cher, Miyamoto, & Lenert, 1997). In general, risk averse utility functions are more favorable towards therapies that confer greater chances for short term survival because such therapies provide higher probabilities of survival during the period in which risk averse utilities increase most rapidly. One of the primary reasons why UA Problem 2 is important is that a solution to this problem yields a characterization of the risk posture of the client with respect to the continuum in question.

Next we will consider assessment problems that arise when utility is based on a quality adjusted life years (QALY) model. The *linear QALY model* assumes that the utility of survival in any fixed health state is risk neutral. Furthermore, it represents the utility of alternative health states as factors that multiply the duration of survival. In other words, let a pair, (b, x), represent a survival of x years in health state b. The linear QALY model asserts that

$$U(b, x) \;=\; k \cdot H(b) \cdot x. \qquad (1)$$

The left panel of Figure 2 shows the linear QALY model with three health states, a, b, and c. The function H maps health states to health state utilities. It is standard practice to assign the value, $H($full health$) = 1$. Thus, state a represents full health in Figure 2. All other better-than-death health states are assigned utilities between 1 and 0. The constant k is an arbitrary constant chosen so that the utilities range over a convenient interval of numbers. For example, if the longest survival to be considered in the decision analysis is 25 years, one can set $k = 4$. Under this choice, U(full health, 25 years) = 4 · H(full health) · 25 = 100. Since the utility of 0 years is 0, this choice of k yields utilities that range between 0 and 100.

Equation (1) describes the utility of a survival of x years in a constant (chronic) health state b. To apply the linear QALY model to a sequence of health states that change over time, one assumes that different time periods contribute additively to the overall utility . In other words, let $(b_1, x_1; ....; b_n, x_n)$ stand for a health sequence where health state b_1 lasts for x_1 years, ..., and health state b_n lasts for x_n years, followed by death (any of the durations x_i can be a fraction of a year, if necessary). The linear QALY model applied to this health sequence asserts that

$$U(b_1, x_1; ....; b_n, x_n) = \Sigma\, k \cdot H(b_i) \cdot x_i. \qquad (2)$$

To illustrate this equation, suppose that we wish to calculate the utility of a sequence $(b_3$, *3 years*; b_1, *12 years*; b_2, *5 years*; b_3, *5 years*$)$. Equation (2) states that

$$U(b_3, 3\ years;\ b_1, 12\ years;\ b_2, 5\ years;\ b_3, 5\ years)$$

$$= k \cdot H(b_3) \cdot 3 + k \cdot H(b_1) \cdot 12 + k \cdot H(b_2) \cdot 5 + k \cdot H(b_3) \cdot 5. \qquad (3)$$

The left panel of Figure 3 shows the implications of the additivity assumption (2) for the hypothetical case (3). The bottom left graph shows linear utility functions for survival in health states b_1, b_2, and b_3, assuming these health states to be constant. The upper left graph shows how the segments of the constant health state utility functions are combined according to Equation (2) to yield the utility of the sequence. The total utility of the sequence is indicated by the height of the point marked U.

The key assessment problem for the linear QALY model is the determination of the health state utilities, *H(a)*, *H(b)*, *H(c)*, ...

UA Problem 3. Assuming the validity of the linear QALY model (1), determine the health state utilities, *H(b)*, for the various health states b in the decision analysis.

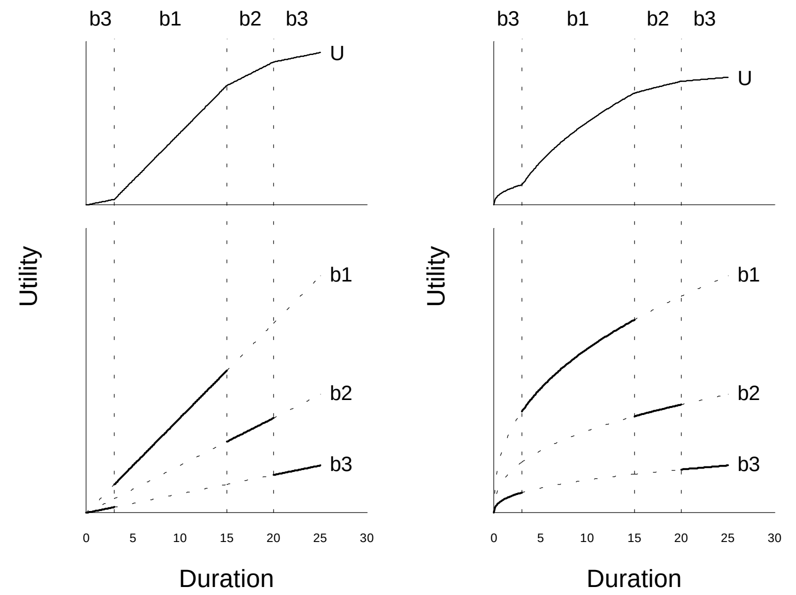

Figure 3. Left panel: Additivity across time periods in a linear QALY model. Right panel: Additivity across time periods in a power QALY model.

Given the health state utilities of a client, Equation (1) can be used to calculate utilities for chronic health states, and Equation (2) can be used to calculate utilities for sequences of health states. Most discussions of health utility analysis refer to the linear QALY model as simply *the QALY model*. In other words, when publications refer to *the* QALY model, it is assumed that Equation (1) and (2) describe the calculation of utility. This chapter uses the nonstandard term, "linear QALY model," because an alternative, power QALY model will also be discussed.

One limitation of the linear QALY model is that it assumes that individuals are risk neutral with respect to survival duration. The following *power QALY model*[1] allows for risk aversion or risk seeking.

$$U(b, x) = k \cdot H(b) \cdot x^r. \quad (4)$$

The power QALY model implies that a utility function for survival duration is risk averse if $r < 1$, it is risk neutral if $r = 1$, and it is risk seeking if $r > 1$. The three utility functions shown in Figure 1 are power utility functions with r set at .5, 1, and 2, respectively. The right panel of Figure 2 shows the power QALY model with three health states, *a*, *b*, and *c*, for the specific power $r = .5$. As before, k is an arbitrary constant. If *H(full health)* is set equal to 1 and if 25 years is the longest duration in the decision analysis, then choosing $k = 100/25^r$ causes the utilities to range between 0 and 100.

The power QALY model (4) describes the utility of a survival of x years in a constant health state b. To calculate the utility of a sequence of health states, one assumes that different time periods contribute additively to the overall utility. For any sequence, $(b_1, x_1; ....; b_n, x_n)$, let $x_0 = 0$. Then,

$$U\left[(b_1, x_1), \ldots, (b_n, x_n)\right] = \sum_{i=1}^{n} k \cdot H(b_i) \cdot \left[\left(\sum_{k=0}^{i} x_k\right)^r - \left(\sum_{m=0}^{i-1} x_m\right)^r\right]. \quad (5)$$

Although Equation (5) may look complicated, the intuition behind it is identical to the additivity assumption for the linear QALY model. The lower right panel of Figure 3 shows power utility functions for survival duration in constant health states b_1, b_2, and b_3. To compute the utility of the sequence, *(b_3, 3 years; b_1, 12 years; b_2, 5 years; b_3, 5 years)*, one takes the corresponding segments from the lower right panel, and pieces them together to form the utility function in the upper right panel. The total utility of the sequence is indicated by the height of the point marked U. Equation (5) is simply an algebraic description of this construction.

1 More precisely, one might call Model (4) the multiplicative power model. Miyamoto (in press) discusses more general versions of the power model.

The main assessment problem for the power QALY model is the assessment of r and the values of H.

UA Problem 4. Assuming the validity of the power QALY model (4), determine the value of the power parameter r, and the health state utilities, *H(b)*, for the various health states b in the decision analysis

Given estimates of r and the values of H for a particular client, one can use Equation (4) to model utility in constant or chronic health states, and Equation (5) to model the utility of sequences of health states[2].

A brief word on the history of these models. The linear QALY model has a lengthy history that is recounted in Fryback (in press) and Drummond et al. (1997). Axiomatic work on QALY models began with Pliskin, Shepard, and Weinstein (1980). They published a set of axioms for the linear QALY model (1), and the more general, power QALY model (4) under EU assumptions. They showed that the linear QALY model is valid if preferences satisfy four properties: (i) survival duration is utility independent from health quality; (ii) health quality is utility independent from survival duration; (iii) proportional time tradeoffs are constant; and (iv) preferences for lotteries over survival duration are risk neutral. They further showed that the power QALY model is valid if assumptions (i), (ii), and (iii) are satisfied, , (iv)and assumption (iv) is replaced with the assumption that marginality is violated. Bleichrodt, Wakker, & Johannesson (1997) pointed out that the axioms for the linear QALY model could be substantially simplified if one assumes that different health states are equally preferred when the survival duration is zero. This assumption was called the zero condition by Miyamoto, Wakker, Bleichrodt, and Peters (in press), who review its history and derive further implications from it. Bleichrodt, Wakker and Johannesson showed that the zero condition and risk neutrality are jointly sufficient for the linear QALY model (1). Miyamoto (in preparation) showed that the zero condition and constant proportional risk posture are jointly sufficient for the power

2 Cher, Miyamoto, and Lenert (1997) explain an alternative way to compute the utility of a health sequence for a power QALY model. Their method uses derivatives of the utility function, and yields an approximation to the utility of the sequence. Equation (5) yields an exact value, assuming the validity of a power QALY model.

QALY model (4). Both results are special cases of a general theorem in Miyamoto (1992) that showed that the zero condition and the utility independence of survival duration are jointly sufficient for a model in which the utility of duration and health quality combine multiplicatively and converge at zero duration. Miyamoto and Eraker (1988) investigated the empirical validity of the utility independence of survival duration, and Miyamoto and Eraker (1989) investigated the empirical validity of risk neutrality and constant proportional risk posture (a necessary condition for a power QALY model).

Utility Assessment Under Expected Utility (EU) Assumptions

Basic Notation

To discuss utility assessments, we will need some notations for lotteries and preferences.

Notation	*What it stands for:*
A, B, C, ...	Health outcomes.
(A, p; B, 1-p)	A lottery in which one has a *p*-chance of receiving health outcome *A* and a 1-*p* chance of receiving health outcome *B*.
(A, p; B, 1-p) $\succ$ *(C, q; D, 1-q)*	*(A, p; B, 1-p)* is preferred to *(C, q; D, 1-q)*
(A, p; B, 1-p) $\sim$ *(C, q; D, 1-q)*	*(A, p; B, 1-p)* and *(C, q; D, 1-q)* are equally preferred.
(A, p; B, 1-p) $\succeq$ *(C, q; D, 1-q)*	*(A, p; B, 1-p)* is equally or more preferred than *(C, q; D, 1-q)*

For example, if *A* represents "full health" and *B* represents a specific inferior health state, then *(A, .75; B, .25)* represents a lottery in which one has a 75% chance of full health and a 25% chance of the inferior health state.

Basic EU Theory

EU theory is a theory of preference under risk. When applied in a health domain, the basic objects of EU theory are lotteries for health outcomes and the fundamental empirical relation is the preference relation among such lotteries. Although health outcomes can be complex sequences of health states unfolding over time, and lotteries can also be complex, the outcomes and lotteries that are used in utility assessment are only the most elementary types. The only options required for utility assessment are simple outcomes (riskless outcomes) and binary lotteries (lotteries with two outcomes). The basic claim of EU theory is that the preference ordering among lotteries is the same as the ordering of the lotteries by their expected utilities. This claim can be stated in terms of binary lotteries as follows:

$$(A, p; B, 1\text{-}p) \succ (C, q; D, 1\text{-}q)$$

$$\text{iff}$$

$$pU(A) + (1\text{-}p)U(b) > qU(C) + (1\text{-}q)U(D) \qquad (6)$$

and

$$(A, p; B, 1\text{-}p) \sim (C, q; D, 1\text{-}q)$$

$$\text{iff}$$

$$pU(A) + (1\text{-}p)U(b) = qU(C) + (1\text{-}q)U(D). \qquad (7)$$

(The expression "iff" is an abbreviation for "if and only if.") Condition (6) states that one gamble is preferred to another if and only if its expected utility is greater. Condition (7) states that equivalence in preference maps onto equality of expected utility. Conditions (6) and (7) represent the hypothesis of EU maximization for the special case of binary lotteries. For lotteries with more than two outcomes, one postulates that the EU of a lottery equals $\Sigma\, p_i U(x_i)$ where p_i represents the probability of receiving outcome x_i. Such more complicated notations will not be required in this chapter.

One special case of Condition (7) is especially useful in utility assessments, namely, the case where one observes an equivalence between a certain outcome *C* and a gamble *(A, p; B, 1-p)*. From Condition (7), we may infer that

$$C \sim (A, p; B, 1\text{-}p) \text{ iff } U(C) = pU(A) + (1\text{-}p)U(b). \qquad (8)$$

Many utility assessment procedures require that one observe equivalences like the left side of (8) and then utilities are inferred from numerical relationships that are implied by the right side of (8).

For purposes of utility assessment, it is also important that if a utility function U satisfies the EU assumptions, then U is a *cardinal utility* or equivalently, U is an *interval scale*. A precise definition of these terms requires the use of set theory (see Krantz, Luce, Suppes, & Tversky, 1971, or Roberts, 1979), but the essential idea is that in measuring a utility function, one is allowed two arbitrary assignments of utility. To give an analogy, if one were asked to assign coordinates to the points on an infinite straight line (infinite in both directions), one could pick any point and call it zero, and pick any other point and call it a unit distance from zero. After these two arbitrary choices, the coordinates of all other points would have definite values. In the context of utility assessment, researchers usually choose to assign 0 to the utility of death, and 100 to the utility of full health. The logic of utility assessment and the empirically determined preferences of an individual then force all other outcomes to be assigned specific numerical utilities.

Conditions (6) **-** (8) exhibit a key property of EU theory. According to EU theory, utility is linear in probability. In other words, the probabilities by which the lotteries are defined, e.g., *p* and *q* on the left side of (6) or (7), are used directly in the calculation of the expected utility, e.g., *p* and *q* also appear on the right side of (6) or (7). Contrast this with the hypothesis that the probabilities are transformed nonlinearly in the cognitive process by which a risky option is evaluated. For example, suppose that in place of Condition (6), we had the condition:

$$(A, p; B, 1\text{-}p) \succ (C, q; D, 1\text{-}q)$$

$$\text{iff}$$

$$w(p)U(A) + \big(\mathbf{1} - w(p)\big)U(b) > w(q)U(C) + \big(\mathbf{1} - w(q)\big)U(D) \qquad (9)$$

In Condition (9), a nonlinear function *w* transforms the probabilities, *p* and *q*, to decision weights (psychological weights). In later sections, we will explore the implications of nonlinear probability weighting for the methodology of utility assessment. For now, I want to draw attention to the fact that EU theory implies that utility is linear in probability. This assumption plays a central role in the EU theory of utility assessment.

Conditions (6) **-** (8) express all of the formal part of EU theory that is needed to discuss utility assessment procedures, but there is, of course, a great deal more that is relevant to the validity of utility assessments. Underlying EU theory is a set of preference assumptions, known as EU axioms, from which Conditions (6) **-** (8) and other related conditions can be derived. It should be noted that calling an assumption an "axiom" does not imply the empirical validity of the assumption. An axiom is simply one of a set of assumptions that are jointly sufficient to imply the validity of a theory, in this case, EU theory. The assumption that preferences are transitive, or the betweenness assumption (if $A \succ B$ and $1 > p > 0$, then $A \succ (A, p; B, 1\text{-}p) \succ B$), are examples of EU axioms. It can then be proved that if preferences are consistent with the EU axioms[3], then there exists a utility function U that satisfies Conditions (6) **-** (8) and other related conditions. In this chapter, I will use the expression, "EU assumptions," to refer to preference assumptions that are either EU axioms or are implied by the EU axioms. To assert that EU theory is descriptively valid is to assert that all EU assumptions are empirically valid properties of preference behavior. Conversely, when researchers claim that EU theory is descriptively invalid, they mean that at least some of the EU assumptions are violated by actual preference data. In fact, there is a great deal of evidence that preferences are inconsistent with the assumptions of EU theory. I will not attempt to review empirical tests of EU assumptions (see Kahneman & Tversky, 1979, 1984; Slovic et al., 1988; Camerer, 1989; Luce, 1992), and will assume that the evidence against the descriptive validity of EU assumptions is quite strong.

The EU Theory of Assessment Procedures

Most of the methods presented in this section are all well known, but they are reviewed here in order to have an explicit point of comparison

[3] To be precise, there is not one unique set of EU axioms. Rather, theoretical analyses have uncovered a variety of alternative sets of assumptions any one of which is sufficient to imply the validity of EU theory. Any of these sets of assumptions can be called *a* set of axioms for EU theory. Fishburn (1982) reviews alternative EU axiomatizations.

Table 1

Duration	EU Assessments		RDU Assessments	
	p^*	$(23.8)X^{.446}$	$w(p^*)$	$(13.67)X^{.618}$
0	0.00	0.0	0.000	0.0
5	0.45	48.8	0.395	37.0
10	0.65	66.5	0.503	56.8
15	0.85	79.6	0.654	72.9
20	0.99	90.5	0.912	87.1
25	1.00	100.0	1.000	100.0

with the non-EU approach in the next section. More thorough descriptions of these methods are available in Sox et al. (1988), Froberg and Kaplan (1989b), and Drummond et al. (1997). The emphasis here will be on the logic by which utilities are inferred from preferences, and the role of linearity in drawing these inferences. For the sake of brevity, I will not address statistical issues that arise in utility assessment beyond what is necessary to explain the specific examples in this chapter. The utility assessment methods described in this section all assume the descriptive validity of EU theory. This assumption will be dropped in the subsequent section on utility assessment under rank-dependent utility assumptions.

In the *standard gamble method*, a best and worst outcome are identified for the given health domain. Let A designate the best outcome, and let Z designate the worst outcome. As noted above, we are free to assign the utilities U(A) = 100 and U(Z) = 0. To assess the utility of any other outcome, B, the client is asked to judge the probability *p** that satisfies the relation:

$$B \sim (A, p^*; Z, 1-p^*). \qquad (10)$$

If *p** is the probability that creates the equivalence (10), then *p** will be called the *probability equivalent* of B with respect to the endpoints A and Z[4]. By Condition (8), we infer that

$$U(B) = p^*U(A) + (1-p^*)U(Z) = p^*(100). \qquad (11)$$

For example, if the client says that a *.8* chance of *A* and a *.2* chance of *Z* is equal in preference to *B*, then *U(B) = 80*. Clearly, the standard gamble zmethod provides a straightforward solution to UA Problem 1.

The standard gamble method also provides solutions to the remaining three assessment problems. To solve UA Problem 2, let *Z* denote 0 years and let *A* denote the longest survival duration in the assessment problem. Assume that health state is fixed at some better-than-death health state. One then applies (10) and (11) to determine the utilities of a series of intermediate points, $X_1, ..., X_n$. Linear interpolation between these points provides a piecewise linear utility function that approximates the utility function for survival duration. To illustrate this procedure, column 2 of Table 1 shows hypothetical probability equivalents for the durations 5, 10, 15, and 20 years with respect to the endpoints 0 and 25 years. Multiplying these probability equivalents by 100 yields utilities for the durations scaled from 0 to 100. The solid lines in the left panel of Figure 4 show the piecewise linear utility function for the data in Column 2 of Table 1.

As an alternative to the piecewise linear utility function, one could fit a parametric utility function like a power or exponential utility function to the pairs, $[X_1, U(X_1)], ..., [X_n, U(X_n)]$, by means of a nonlinear regression procedure. Such procedures are available, for example, in the S-Plus, SPSS, and SAS statistical packages. To illustrate this approach, let $X_1, ..., X_n$ be a list of survival durations, and let $Z = 0$ and *A* denote the worst and the best survival durations. Let $p_1^*, ..., p_n^*$ denote the probability equivalents of $X_1, ..., X_n$ with respect to the endpoints, *Z* and *A*, and let $U(X_1), ..., U(X_n)$, be the corresponding utilities inferred by means of the standard gamble method. According to the power QALY model, $U(X_i) =$

4 Some authors refer to p* as the *indifference probability* of B (cf., Sox et al., 1988).

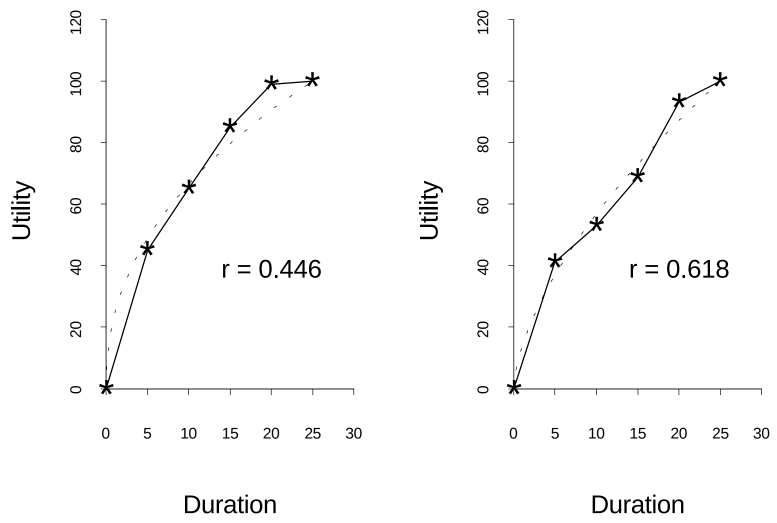

Figure 4. Left panel: Utilities assessed by the method of standard gambles under EU assumptions. The solid line is the piecewise linear approximation to the utility function; the dotted line is the power function approximation. Right panel: Utilities assessed by transforming probability equivalences to probability weights under RDU assumptions. The solid line is the piecewise linear approximation to the utility function; the dotted line is the power function approximation.

$k \cdot X_i^r$ for every i. Because utilities were assigned under the specification $U(A) = 100$, we must have $100 = U(A) = k \cdot A^r$, or $k = 100/A^r$. By Equation (11), $U(X_i) = p_i^*(100)$, where p_i^* is the i-th probability equivalent. Therefore $p_i^*(100) = U(X_i) = (100/A^r) \cdot X_i^r$, i.e.,

$$p_i^* = X_i^r / A^r = (X_i/A)^r \tag{12}$$

To fit a power function to utility data that were assessed by means of standard gambles, one lets the values of p_i^* serve as the dependent variable, and the values of (X_i/A) serve as the predictor variable in a nonlinear regression[5] that solves for the value of r. Applying the nls procedure of S-Plus to the data[6] in columns 1 and 2 of Table 1 yielded a fit of the power utility function, $U(X) = (23.8)X^{.446}$. The dotted line in the left panel of Figure 4 shows the fitted power model.

If one assumes that EU theory is descriptively valid and that the responses to the standard gamble questions are free from random variation, then the standard gamble method yields exact values of the utility function. In a sense, then, to approximate the standard gamble utilities by means of a power function (or any other function) is a step away from accurate measurement because the data themselves are precisely correct utility measurements. Nevertheless, there are two reasons for taking an interest in a parametric utility representation like a power utility function. First, human judgment almost always exhibits random variation in the sense that asking the identical assessment question to the same client will produce somewhat different responses at different times. Even if EU theory were descriptively valid, utilities assessed by the standard gamble method, or any other method, for that matter, would not be precisely accurate because they are affected by random variation in judgment. Fitting a parametric utility function to a set of data is one way to reduce the influence of random variation by aggregating across responses. From this perspective, the fitted curve in Figure 4 (left panel) is a more accurate representation of preference than are the individual data points, because the individual points are subject to greater random variation than a summary constructed from the data.

Second, even if the assessed standard gamble utilities represent true preferences, i.e., even if they were not affected in part by random variation in judgment, the fitting of a power parameter facilitates comparisons of risk posture across individuals, across populations, and across decision analyses. For example, Miyamoto and Eraker (1985) fitted

5 A power utility function can also be fit by means of linear regression, i.e., if $p^* = (X/A)^r$, a logarithmic transformation yields $\log p^* = r \cdot \log(X/A)$. Therefore one can use linear regression through the origin to solve for r. Miyamoto and Eraker (1985) used a method similar to this to estimate r for individual patients.

6 It was necessary to omit the initial point, (0, 0), from the data because the utility function is discontinuous at this point.

the power utility model to the certainty equivalents data of individual patients. They found that estimates of r were about equally divided between values greater than and less than 1. In other words, risk aversion and risk seeking were found about equally often in their data. The power parameter allows us to state this finding even if the power utility model is not a precisely accurate description of the utilities. To give another example, in the next section we will reinterpret the data in Table 1 from the standpoint of rank-dependent (RDU) utility theory. The analysis will show that under RDU assumptions, the estimated power parameter is .618 rather than .446 as found under EU assumptions. This shows that the RDU analysis yields a utility function that is less risk averse than the EU analysis. The fitting of a power parameter allows one to state concisely an interesting relationship between different utility functions. This issue will be discussed further below.

To solve UA Problem 3, assume the descriptive validity of EU theory and the linear QALY model. Let *a* represent full health, and set *H(a)* = 1. Suppose we want to determine H(b) for a better-than-death health state b. Choose any survival duration y and use the standard gamble method to find the probability *p** that the client judges to satisfy the equivalence,

$$(\mathrm{b}, \mathrm{y}) \sim \left[(\mathrm{a},\mathrm{y}), \mathrm{p}^*; (\mathrm{a},\mathbf{0}), \mathbf{1}-\mathrm{p}^*\right]. \qquad (13)$$

By Equation (8) and the linear QALY model (1), we have

$$k \cdot H(b) \cdot y = p^* \cdot k \cdot H(a) \cdot y + (1\text{-}p^*) \cdot k \cdot H(a) \cdot 0 \qquad (14)$$

Thus, $$H(b) = p^*, \qquad (15)$$

because *H(a)* = 1. Evidently, one can use this procedure repeatedly to find the values of H(b) for any finite list of better-than-death health states. Thus, the standard gamble method solves UA Problem 3 under EU assumptions.

To solve UA Problem 4, note first that the assessment of *H(b)* by means of standard gambles is the same for a power QALY model as for a linear QALY model. To see this, suppose that *p** satisfies (13). By Equation (8) and the power QALY model (4),

$$k \cdot H(b) \cdot y^r = p^* \cdot k \cdot H(a) \cdot y^r + (1\text{-}p^*) \cdot k \cdot H(a) \cdot 0^r \qquad (16)$$

Thus, $$H(b) = p^*. \qquad (17)$$

Therefore the empirical relation, (13), determines the health state utility by means of the identical Equations (15) or (17) for either the linear or the power QALY models. To complete the assessment of the power QALY model, one needs to assess *r*. The solution to this assessment problem was sketched above. Choose any fixed better-than-death health state *b*. Often, one would choose *b* equal to either current symptoms or to the best health state in the decision analysis, but from the standpoint of logic, any choice of *b* is permissible. Let 0 and *A* be the shortest and longest survival durations, respectively. Let (b, X_1), ..., (b, X_n) be a list of intermediate outcomes, and let $\mathrm{p}^*_{\mathbf{1}}$, ..., $\mathrm{p}^*_{\mathrm{n}}$ be the corresponding probability equivalents with respect to the endpoints, *(b, 0)* and *(b, A)*. According to Equations (4),

$$k \cdot H(b) \cdot \mathrm{X}_\mathrm{i}^\mathrm{r} = \mathrm{p}_\mathrm{i}^* \cdot k \cdot H(b) \cdot A^r + (1-\mathrm{p}_\mathrm{i}^*) \cdot k \cdot H(b) \cdot 0^r \qquad (18)$$

where (18) follows from (4) and (11). Therefore

$$\mathrm{p}_\mathrm{i}^* = (\mathrm{X}_\mathrm{i}/\mathrm{A})^\mathrm{r} \qquad (19)$$

exactly as was found before. An estimate of *r* can then be determined by nonlinear regression with $\mathrm{p}^*_{\mathbf{1}}$, ..., $\mathrm{p}^*_{\mathrm{n}}$ as the values of the dependent variable and X_1/A, ..., X_n/A as the values of the predictor variable. Thus, the standard gamble method yields a solution to UA Problem 4 if one assumes EU theory and the power QALY model.

Notice that in all of the applications of the standard gamble method, the validity of the assessment is heavily dependent on the assumption that utility is linear in probability. For each utility that is to be assessed, the client must produce a *p** that satisfies an equivalence of the form of (10) or (13), and the calculation of the utility requires that *p** and 1–*p** constitute the appropriate weights for the superior and inferior utilities, respectively, on the right side of Equation (11). If the perception of probability is systematically distorted, the standard gamble method transfers the distortions into the utility scale by means of Equation (11).

This completes our discussion of the standard gamble method under EU assumptions. Next we will consider the time tradeoff procedure, and finally, the method of certainty equivalents.

The time tradeoff method was introduced by Torrance, Thomas, and Sackett (1972) as a measure of health status. Let *a* denote full health or whatever is the best health state in the utility assessment problem, and let b be any other better-than-death health state. Let $x > 0$ be any survival duration. Then the *time tradeoff between a and b with respect to the duration x* is the duration y^* such that

$$(a, y^*) \sim (b, x). \tag{20}$$

If y^* satisfies (20), then *the proportional time tradeoff between a and b with respect to duration x* is the ratio y^*/x. For brevity, I will refer to time tradeoffs as TTOs and proportional time tradeoffs as PTTOs. The relation between TTOs and health state utilities depends on the utility assumptions that one adopts. If one assumes only the validity of EU theory, but not the validity of the linear or power QALY models, then (20) implies only that $U(a, y^*) = U(b, x)$, and nothing more.

If one assumes the validity of EU theory and the linear QALY model (1), then the TTO method provides a solution to UA Problem 3. Let a be the best health state, and assign $H(a) = 1$. For any better-than-death health state b, determine the y^* that satisfies (20). Then (20) implies that

$$k \cdot H(a) \cdot y^* = k \cdot H(b) \cdot x, \tag{21}$$

hence,

$$H(b) = y^*/x. \tag{22}$$

because $H(a) = 1$. Therefore, assuming EU theory and the linear QALY model, TTOs with respect to the best health state provide a solution to UA Problem 3. TTOs also contribute to the solution of UA Problem 4, but to explain this, one must first explain the use of certainty equivalents to assess a power parameter.

The method of certainty equivalents is used primarily to solve UA Problems 2 and 4. Let (b, x) and (b, z) denote any two survival durations in a constant health state b. Then, we say that (b, y^*) is the certainty equivalent of the lottery, $[(b,x), p; (b,z), \mathbf{1}-p]$, if and only if

$$(b, y^*) \sim [(b,x), p; (b,z), \mathbf{1}-p]. \tag{23}$$

In this chapter, we will only be concerned with certainty equivalents of even-chance gambles, i.e., lotteries of the form $[(b,x), 0.5; (b,z), 0.5]$. Even-chance gambles are especially useful in utility assessment because the concept of a flip of a fair coin is widely understood by the general public. Suppose that (b, y^*) is the certainty equivalent of the even-chance gamble, $[(b,x), 0.5; (b,z), 0.5]$, i.e.,

$$(b, y^*) \sim [(b,x), 0.5; (b,z), 0.5]. \tag{24}$$

Under EU assumptions:

$$\hat{y}^* = [.5 \cdot x^{.74} + .5 \cdot z^{.74}]^{1/.74}$$

Under RDU Assumptions

$$\hat{y}^* = [.44 \cdot x^{.90} + .56 \cdot z^{.90}]^{1/.90}$$

Equation (8) and the power QALY model imply that

$$k \cdot H(b) \cdot (y^*)^r = .5 \cdot k \cdot H(b) \cdot x^r + .5 \cdot k \cdot H(b) \cdot z^r \tag{25}$$

Therefore

$$(y^*)^r = .5 \cdot x^r + .5 \cdot z^r \tag{26}$$

and

$$y^* = [.5 \cdot x^r + .5 \cdot z^r]^{1/r}. \tag{27}$$

Once again we have a problem in nonlinear estimation. To estimate r, collect data for certainty equivalents with varying values of x and z. Let the certainty equivalents serve as the dependent variable, and the values of x and z serve as the predictor variables in a nonlinear regression that solves for the value of r. To illustrate this idea, Table 2 contains hypothetical certainty equivalents data for five even-chance gambles. The data appear to be slightly risk averse, and the fit of Equation (27) to these data yields an estimate of $r = .74$ which is slightly less than 1. Of course, in actual

Table 2
Hypothetical Certainty Equivalents
$y^* \sim (x, .5; z, .5)$

y^*	x	z
4	10	0
6	12	2
10	25	0
12	24	2
12	24	4

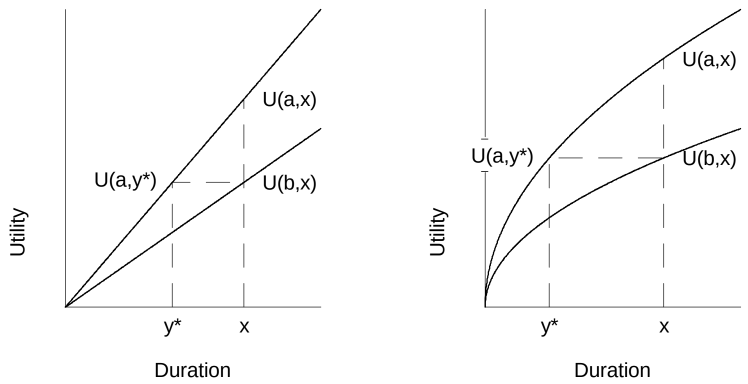

Figure 5. Left panel: In the linear QALY model, $U(b, x)/U(a, x) = y^*/x$. Right panel: In a power QALY with a risk averse utility function, $U(b, x)/U(a, x) > y^*/x$.

research, one would prefer to have more data. This solves UA Problem 2 by means of certainty equivalents.

To solve UA Problem 4 by means of certainty equivalents, we also need to consider TTOs. Let *a* be the best health state so that $H(a) = 1$. Assuming the validity of EU theory and the power QALY model (4), the TTO shown in (20) implies that

$$k \cdot H(a) \cdot (y^*)^r = k \cdot H(b) \cdot x^r, \tag{28}$$

hence

$$H(b) = \left(\frac{y^*}{x}\right)^r. \tag{29}$$

Notice that EU theory and the linear QALY model imply that (22) gives the value of *H(b)*, whereas EU theory and the power QALY model imply that (29) gives the value of *H(b)*. Because the parameter *r* represents risk posture in the power QALY model, Equation (29) is sometimes said to define a *risk-adjusted PTTO*. I will refer to an estimate calculated by means of (29) as an *RA-PTTO*.

To gain some intuition for the role of *r* in Equation (29), consider Figure 5. If the client judges that the equivalence (20) holds, $U(a, y^*) = U(b, x)$. The linear QALY model and the power QALY model both imply that

$$\frac{U(b,x)}{U(a,x)} = \frac{H(b)}{H(a)} = H(b). \tag{30}$$

The ratio, $U(b,x)/U(a,x)$, is not directly observable; rather, what one can observe is the ratio, y^*/x. If the utility functions for survival duration are straight lines radiating from the origin as shown in the left panel of Figure 5, then $H(b) = U(b, x)/U(a, x) = y^*/x$. As shown in the right panel of Figure 5, the curvature of a risk averse utility function causes the ratio y^*/x to exaggerate the difference between health states *a* and *b*. Thus, $H(b) = U(b, x)/U(a, x) > y^*/x$. It is not hard to see that if the utility function is risk seeking, $U(b, x)/U(a, x)$ is less than y^*/x. Therefore, raising the PTTO to the power *r*, as shown in Equation (24) is a correction for the curvature of the utility functions. It removes a distortion in the assessment of $H(b) = U(b, x)/U(a, x)$ that is introduced by the risk posture of the utility function. The PTTO, y^*/x, overstates the reduction in utility if the utility of survival duration is risk averse, and it understates the reduction if the utility of survival duration is risk seeking.

Certainty equivalents and TTOs provide another solution to UA Problem 4. Certainty equivalents and Equation (27) allow one to estimate the power parameter *r*. Then, TTOs with respect to the best health state and Equation (29) allow one to estimate *H(b)* for all other health states.

The standard gamble method was introduced by the creators of EU theory, von Neumann and Morgenstern (1944). Torrance et al. (1972) described the measurement of health state utilities by means of a variant of the standard gamble procedure and introduced the TTO method as an alternative method for determining health state utilities in a linear QALY model. Torrance (1986) presented a thorough review of the theoretical foundations of health utility assessment under EU assumptions (for more recent reviews, see Froberg & Kaplan, 1989a, 1989b, 1989c, 1989d; and Drummond et al., 1997). Miyamoto and Eraker (1985) presented a method for assessing the power parameter of the power QALY model (4) from certainty equivalents data. The method presented in this chapter is an improvement over this method. They also emphasized the need to adjust TTOs for risk posture when assessing health state utilities for a power

QALY model. Cher, Miyamoto, and Lenert (1997) pointed out that Markov process models of clinical decisions can be sensitive to the distinction between TTO and RA-TTO measures of health state utilities.

Summary of EU Assessment Methods

I will summarize the assessment procedures that are sufficient to solve UA Problems 1 - 4 under EU assumptions.

UA Problem 1. Assess the utilities of holistic outcomes.

- The standard gamble method provides the only solution when no attribute structure is specified for the outcomes.

UA Problem 2. Assess the utility of survival duration in some fixed health state.

- The standard gamble method can be used to provide a piecewise linear approximation to a utility function for survival duration. One can also fit a power function to the assessed utilities by means of a nonlinear regression procedure.
- Alternatively, certainty equivalents can be collected, and a power function can be fit to these equivalents by means of a nonlinear regression procedure.

UA Problem 3. Assess the values of H(b), assuming the validity of the linear QALY model.

- The standard gamble method yields values of *H(b)* for each health state *b*.
- Alternatively, PTTOs also yield values of *H(b)* for each health state *b*.

UA Problem 4. Assess the values of r and H(b), assuming the validity of the power QALY model (4):

- The standard gamble method yields utilities for individual survival durations. A power function can be fit to the assessed utilities as in UA Problem 2. The standard gamble method also yields values of *H(b)* for each health state *b*.
- Alternatively, certainty equivalents can be collected, and a power function fit to these certainty equivalents as in UA Problem 2. PTTOs must be adjusted for risk posture as in Equation (29) to yield values of *H(b)* for each health state *b*.

The next section presents solutions to these same assessment problems under rank dependent utility theory.

Rank-Dependent Utility (RDU) Theory

Rank-dependent utility (RDU) theory is a major attempt to explain the violations of EU theory by means of a postulated nonlinear transformation of probabilities. A full discussion of RDU theory requires a description of the representation of lotteries as cumulative probability distributions over outcomes, and of the process by which a nonlinear transformation of cumulative probabilities is converted to decision weights (Quiggin, 1982, 1993; Quiggin & Wakker, 1994). Fortunately, the only lotteries required for the utility assessments of this chapter are binary lotteries. Therefore to explain how these assessments are interpreted under RDU assumptions, it will suffice to describe the RDU representation of binary lotteries and simple outcomes.

For the special case of binary lotteries, RDU theory asserts that the utility of a lottery is determined by the following formula:

$$\mathrm{U}(\mathrm{A}, \mathrm{p}; \mathrm{B}, \mathbf{1}-\mathrm{p}) = \begin{cases} \mathrm{w(p)U(A)} + [1-\mathrm{w(p)}]\mathrm{U(B)} & \text{if } A \succeq B \quad (32) \\ \\ [1-\mathrm{w}(1-\mathrm{p})]\mathrm{U(A)} + \mathrm{w}(1-\mathrm{p})\mathrm{U(B)} & \text{if } B \succeq A \quad (32) \end{cases}$$

where *w* is a nonlinear function from probabilities to the unit interval, i.e., $1 \geq w(p) \geq 0$ for every probability *p*. It can be shown that under the assumptions of RDU theory the utility function is an interval scale (Wakker, 1989). The utility of a lottery is calculated by Equation (1) when $A \succeq B$, and by Equation (2) when $B \preceq A$. When $A \sim B$, $U(A) = U(B)$, so either (1) or (2) produces the same result. In all analyses discussed here, the lotteries that serve as stimuli have the form *(A, p; B, 1-p)* where *A* is preferred to *B*. Therefore only Equation (1) will be required in the present discussion. Obviously, minor modifications allow one to reformulate the methods with respect to Equation (2) if lotteries are used in which *B* is preferred to *A*.

The exact form of the probability weighting function *w* is presently the subject of intense investigation (Tversky & Kahneman, 1992; Wu & Gonzalez, 1996; Gonzalez, 1993). For purposes of illustration, I will discuss a weighting function suggested by Tversky and Kahneman (1992). Tversky and Kahneman proposed that probability weighting can be represented by the class of transformations:

$$w(p) = \frac{p^{\gamma}}{\left(p^{\gamma} + (1-p)^{\gamma}\right)^{1/\gamma}} \tag{33}$$

For the case of binary lotteries, one can interpret *w(p)* as the weight attached to the utility of the superior outcome[7]. The value of the γ parameter can vary from one individual to the next, corresponding to individual differences in the weight attached to the superior outcome. Figure 6 shows the probability weight *w(p)* when γ equals .61. This value of γ was the median estimate in a sample of 25 Berkeley and Stanford graduate students (Tversky & Kahneman, 1992). The subjects judged certainty equivalents of monetary gambles for gains. In the following discussion, I will assume in some analyses that *w* has the form (33) with $\gamma = .61$. This assumption is made for the sake of illustrating the implications of a specific *w* for utility assessment. More empirical research will be required to determine what are common values of γ to be found in patient populations.

The hypothesis of nonlinear probability weighting alters the theoretical analysis of risk posture. As a concrete example, let us consider the interpretation of the following certainty equivalent under EU and RDU assumptions.

$$\textit{5 years for sure} \sim \textit{(20 years, .5; 0 years, .5)} \tag{34}$$

Panel A of Figure 7 shows the implications of this equivalence under EU assumptions. The utility of 5 years is indicated by an open circle. The height of this circle is halfway between the utility of 0 years

[7] In a full development of RDU theory or CPT, one interprets *w* as a transformation that applies to the cumulative or decumulative probability distribution of a lottery. Discussion of cumulative or decumulative probabilities is unnecessary when one restricts attention to binary lotteries and simple outcomes.

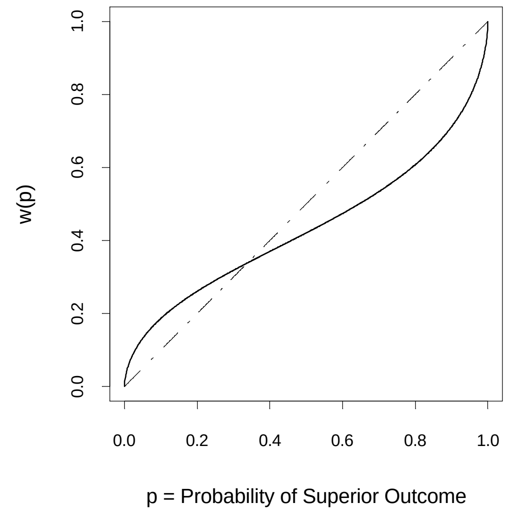

Figure 6. The probability weighting function for gains based on estimates in Tversky and Kahneman (1992).

and the utility of 20 years, as indicated by the dashed lines, because EU theory implies that the utilities of 20 years and 0 years are weighted by the probability, .5. The solid utility curve is a power utility function fit to the datum, *U(5 years) = .5U(20 years)*. The inferred power parameter happens to be .5, i.e., a square root transformation. Panel B depicts an RDU interpretation of the same datum, (34), under the assumption that w(.5) = .4. By Equation (1), *U(5 years) = .4U(20 years)*. The open circle in Panel B is 40% of the height of the utility of 20 years, as shown by the dashed lines on the right of Panel B. Because U(5 years) is lower in Panel B than in Panel A, the power utility function is less risk averse--the inferred power, *r* = .66, is closer to linearity (*r = 1.0*) than the *r* of .50 that was inferred under EU assumptions. Conversely, Panel C shows the implication of the certainty equivalent under the assumption that *w(.5)* = .6. In this case, *U(5 years) = .6U(20 years)*, and the utility function must be more risk averse than the utility function inferred under EU assumptions. Panels D, E, and F show the analogous relations for a risk seeking utility function. If the client judges 15 years for sure ~ (20 years, .5; 0 years, .5), the EU interpretation is that the client is rather risk seeking

(Panel D). If w(.5) < .5, then the RDU interpretation is that the utility function is even more risk seeking (Panel E), and if w(.5) > .5, the RDU interpretation is that the utility function is less risk seeking (Panel F).

Let us say that a probability weighting function w is *optimistic* with respect to p if $w(p) > p$, and it is *pessimistic* with respect to p if $w(p) < p$. Because $w(p)$ is the probability weight for the superior option (see Equation (1)), an optimistic weight places greater weight on the superior option than p, and a pessimistic weight places greater weight on the inferior option than $1–p$. The probability weighting function shown in Figure 6 implies that weights will be optimistic for for $p < .33$, and pessimistic for $p > .33$, where p is the probability of the superior option in a binary lottery. As argued by Wakker and Stiggelbout (1995), the preference behavior that is attributed to risk aversion or risk seeking under EU assumptions may be due in part to nonlinear probability weighting. For example, Panel B of Figure 7 depicts a situation in which a preference that appears to be quite risk averse under EU assumption is attributed under RDU assumptions to a utility function that is somewhat risk averse and a probability weight that is somewhat pessimistic.

Suppose that U_1 and U_2 are two utility functions for survival duration. Let us say that U_1 is *relatively more risk averse* than U_2 if certainty equivalents assuming U_1 are always lower than corresponding certainty equivalents assuming U_2. Converse, let us say that U_1 is *relatively more risk seeking* than U_2 if certainty equivalents assuming U_1 are always greater than corresponding certainty equivalents assuming U_2. In the power utility model, U_1 is relatively more risk averse than U_2 if the power parameter for U_1 is less than the power parameter for U_2 and it is relatively more risk seeking than U_2 if the power parameter for U_1 is greater than the power parameter for U_2. Notice that under this terminology, U_1 can be relatively more risk averse than U_2 even if neither function is risk averse in an absolute sense, i.e., in comparison to a risk neutral utility function. For example, Panels C and F of Figure 7 display utility functions that are relatively more risk averse than Panels B and E, respectively, even though Panel F is not risk averse in an absolute sense.

The essential point from the standpoint of utility assessment is that preference behavior that appears to be risk averse or risk seeking under EU assumptions is interpreted under RDU assumptions to be a consequence jointly of a nonlinear utility function and a nonlinear probability weighting function. Furthermore, if we compare utility functions that are inferred under EU assumptions to utility functions that are inferred under RDU assumptions, the EU functions will be relatively more risk averse than the RDU functions if probability weighting is pessimistic with respect to the probabilities in the assessment, and the EU functions will be relatively more risk seeking than the RDU functions if probability weighting is optimistic with respect to the probabilities in the assessment. The discussion of utility assessment under RDU assumptions will attempt to disentangle the contributions of nonlinear utility and nonlinear probability weighting to the observed preference behavior.

Next I will describe solutions to UA Problems 1 - 4 under RDU assumptions. Equation (1) will be the main analytical assumption. In some assessment methods, it will also be necessary to assume that the weighting function satisfies Equation (33) with $\gamma = .61$. These assumptions will be spelled out as the assesment methods are described. Issues pertaining to risk posture will also be discussed.

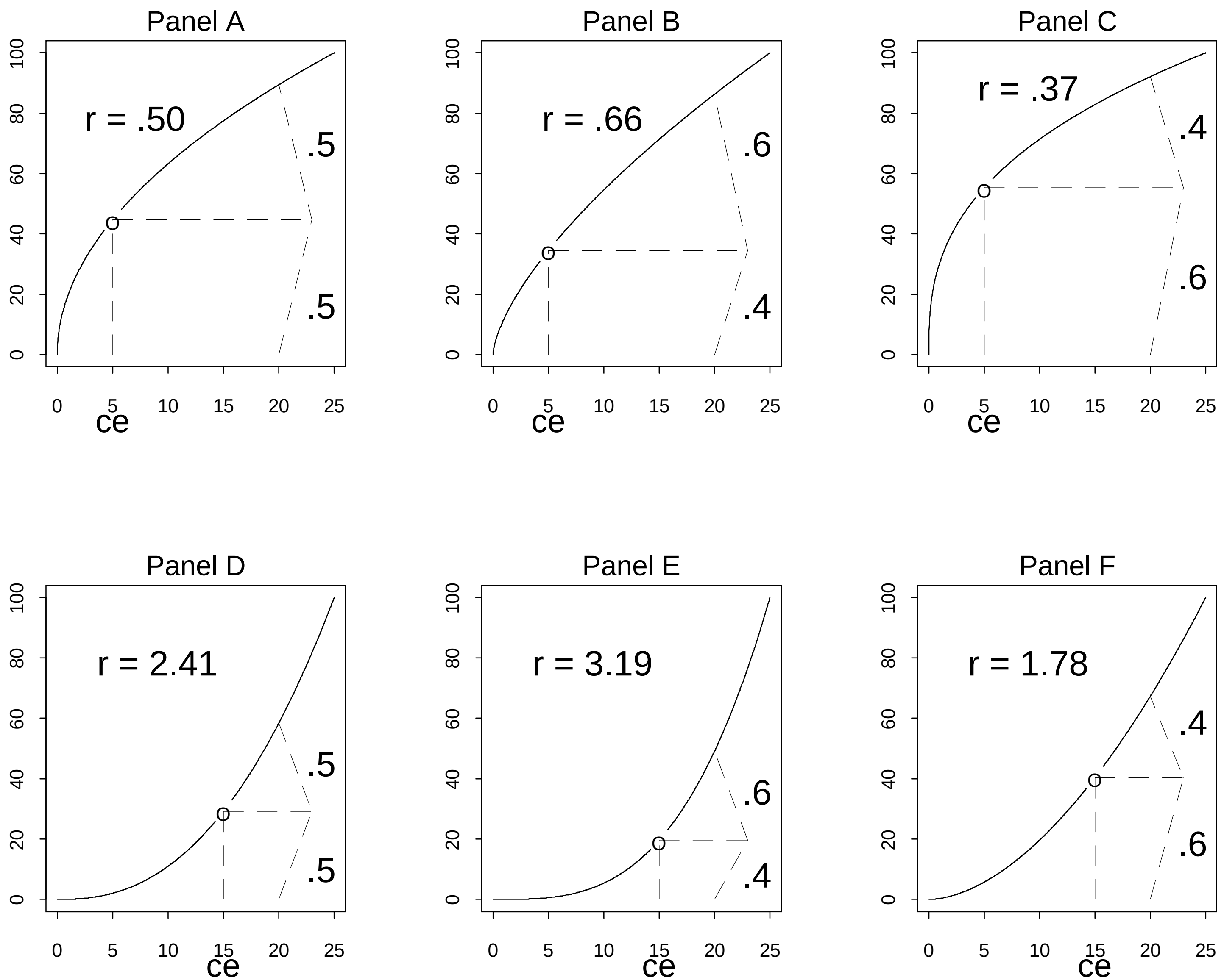

Figure 7. Possible interpretations of a certainty equivalent under EU and RDU assumptions. "ce" indicates the location of the certainty equivalent. Panels A and D assume linear probability weighting. Panels B and E assume pessimistic probability weighting. Panels C and F assume optimistic probability weighting.

Utility Assessment Under RDU Assumptions

In the previous section, it was shown that the standard gamble method provides solutions to UA Problem 1 - 4. The same holds true under RDU assumptions provided that one knows the values of $w(p^*)$ for the probability equivalents that are produced in the assessment. Consider the following argument due to Wakker and Stiggelbout (1995): Let A designate the best outcome, and Z the worst outcome in the health domain under investigation. Let us assign $U(A) = 100$ and $U(Z) = 0$. Let B be any other health outcome, and suppose that p^* is the probability equivalent of B with respect to A and Z as shown in (10). By Equation (1) of RDU theory,

$$U(B) \;=\; w(p^*)U(A) + (1\text{-}w(p^*))U(Z) \;=\; w(p^*)\cdot 100. \qquad (35)$$

Equation (35) can be used to determine the utility of B, provided one knows the value of $w(p^*)$.

Evidently, the key to applying the standard gamble method under RDU assumptions is the development of a method for determining the form of w for individual clients. Although some research articles have published estimates of the weighting function w for groups of subjects (Tversky & Kahneman, 1992; Camerer & Ho, 1994), I do not know of published accounts of estimates of w that are tailored to individuals. I believe, however, that research conditions are ripe for the development of methodologies for assessing the weighting function at the level of individuals. Gonzalez and Wu (unpublished manuscript) have completed an extensive study of alternative models of the probability weighting function, including assessments of model parameters at the level of individuals. Their work is restricted to the study of preferences for monetary lotteries in stimulus designs that were chosen to investigate theoretical questions. Abdellaoui (unpublished manuscript) and Bleichrodt and Pinto (unpublished manuscript) have independently undertaken studies of weighting functions at the individual subject level, the former in the domain of money and the latter in the domain of health. These studies attempt to determine qualitative features of the shape of the probability weighting function without assuming a specific parametric model for probability weighting. So far as I know, it is an open question whether an efficient methodology can be developed for assessing probability weighting functions of individual clients, followed by an RDU assessment of individual health utilities by means of standard gambles. Because assessment of the probability weighting function w holds the key to generalizing the standard gamble method to RDU assumptions and also to cumulative prospect theory assumptions, this would appear to be a productive target for further investigation.

Wakker and Stiggelbout (1995) and Bayoumi and Redelmeier (1996) have applied a group estimate of the probability weighting function to the problem of interpreting standard gamble utilities. A potential deficiency of this approach is that it assumes that all individuals have the same probability weighting function, but it has the virtue of being straightforward to implement at a practical level, and is also heuristically informative. To illustrate this idea, suppose that we wish to assess the utility of an outcome B under the assumption that w has the form of (33) with $\gamma = .61$ as shown in Figure 6. Suppose, further, that the client produces the standard gamble judgment B ~ (A, .8; Z, .2). Applying (33) with $\gamma = .61$, we find that $w(.8) = .607$. By Equation (35), we have $U(B) = 60.7$. Evidently, this process can be repeated for each holistic outcome, thereby providing a solution to UA Problem 1.

To solve UA Problem 2, let $Z = 0$ years, and let A denote the longest survival duration in the assessment problem. Assume that health state is fixed at some better-than-death health state. Let $p_1^*, \ldots, p_n^*$ denote the probability equivalents of a series of intermediate durations, $X_1, \ldots, X_n$. Assuming the validity of Equation (33) with $\gamma = .61$, we can find the utilities of the intermediate durations, $100\cdot w(p_1^*), \ldots, 100\cdot w(p_n^*)$. For example, column 2 of Table 1 shows hypothetical probability equivalents (p^*) for the durations, 5, 10, 15, and 20 years. Column 4 of Table 1 shows the corresponding values of $w(p^*)$ which can be multiplied by 100 to yield utilities scaled from 0 to 100. The right panel of Figure 4 (solid lines) shows a piecewise linear approximation to the utility function for survival duration. The dotted line in the right panel shows a power utility function fitted to these data by means of nonlinear regression as described in the previous section. Notice that the utility function is less risk averse in the RDU analysis--the power parameter r is .618 in the RDU analysis and .446 in the EU. This finding corresponds to the fact $w(p^*) < p^*$ (pessimism) for every probability equivalent in column 2 of Table 1. Whereas EU theory must attribute risk averse preferences entirely to curvature of the utility function, yielding a low $r = .446$, RDU theory can interpret the preferences as a consequence of pessimistic probability weighting and a more moderate curvature of the utility function.

To solve UA Problem 3 under RDU assumptions, assume the descriptive validity of the linear QALY model (1) and Equation (33) with $\gamma = .61$. Let a represent full health, and set $H(a) = 1$. To determine H(b), choose any survival duration y and have the client judge the p* that yields the equivalence:

$$(b, y) \sim [(a, y), p^*; (a, 0), 1 - p^*]. \qquad (36)$$

According to the linear QALY model and Equation (35),

$$k \cdot H(b) \cdot y = w(p^*) \cdot k \cdot H(a) \cdot y + (1 - w(p^*)) \cdot k \cdot H(a) \cdot 0. \qquad (37)$$

Thus, $$H(b) = w(p^*). \qquad (38)$$

For example, if p^* equals .75, then $H(b) = .57$ by Equation (33) with $\gamma = .61$.

To solve UA Problem 4 under RDU assumptions, assume the descriptive validity of the power QALY model (4) and Equation (33) with $\gamma = .61$. Once again, let a represent full health, and set $H(a) = 1$. To determine H(b), have the client judge p^* as in (36). The power QALY model and Equation (35) imply that

$$k \cdot H(b) \cdot y^r = w(p^*) \cdot k \cdot H(a) \cdot y^r + (1 - w(p^*)) \cdot k \cdot H(a) \cdot 0^r. \qquad (39)$$

Thus, $$H(b) = w(p^*) \qquad (40)$$

because $H(a) = 1$. For example, if $p^* = .75$, $H(b) = .57$, exactly as was found for the linear QALY model under RDU assumptions. As can be seen from Equations (38) and (40), under RDU assumptions the standard gamble method yields the same estimate of $H(b)$ for either the linear or power QALY model.

To complete the assessment of the power QALY model under RDU assumptions, let 0 and A be the shortest and longest survival durations, respectively, let $X_1, ..., X_n$ be a list of intermediate durations, and let $p_1^*, ..., p_n^*$ be the probability equivalents of $(b, X_1), ..., (b, X_n)$, respectively, relative to the endpoints (b, A) and $(b, 0)$. Then, the power QALY model and Equation (35) imply that

$$k \cdot H(b) \cdot X_i^r = w(p_i^*) \cdot k \cdot H(b) \cdot A^r + (1 - w(p_i^*)) \cdot k \cdot H(b) \cdot 0^r. \qquad (41)$$

Therefore $$w(p_i^*) = (X_i / A)^r. \qquad (42)$$

Assuming Equation (33) with $\gamma = .61$, we can compute specific values for the $w(p_i^*)$. An estimate of r can then be determined by nonlinear regression with $w(p_1^*), ..., w(p_n^*)$ as the the values of the dependent variable and $X_1/A, ..., X_n/A$ as the values of the predictor variable[8]. Combining this estimate of r with the previous assessment of $H(b)$ for various b yields a solution to UA Problem 4 under RDU assumptions and the assumption of Equation (33) with $\gamma = .61$.

The standard gamble method yields straightforward solutions to UA Problems 1 - 4 if one assumes a specific probability weighting function like Equation (33) with $\gamma = .61$. Clearly, a major question with this procedure is whether individual differences in probability weighting are sufficiently large as to produce substantial deviations from the weighting function with $\gamma = .61$. One should bear in mind that the use here of Equation (33) with $\gamma = .61$ is simply for purposes of illustration. This value of γ was estimated from the preferences of 25 Stanford and Berkeley graduate students for monetary lotteries, and should not be taken too seriously in the context of health utility analysis. What is required are estimates of probability weighting functions that are determined from preferences for health lotteries in populations that are relevant to health utility analysis, e.g., among patients or the general public. Such data would permit one to evaluate whether a single choice of weighting function is a poor or good approximation to the weighting functions of individuals.

The TTO procedure provides a solution to UA Problem 3 under RDU assumptions for precisely the same reasons as under EU assumptions.

---

8 As before, an alternative method of estimation is to treat $\log w(p_1^*)$, ..., $\log w(p_n^*)$ as the values of the dependent variable and $\log(X_1 / A)$, ..., $\log(X_n / A)$ as the values of the predictor variable in a linear regression.

If *(a, y*) ~ (b, x)* as shown in (20), then (21) and (22) follow under RDU assumptions. Therefore $H(b) = y^*/x$, as shown in (22).

The method of certainty equivalents provides alternative solutions to UA Problems 2 and 4. Under RDU assumptions, however, the estimation of r requires that one also estimate $w(.5)$. Let b be any better-than-death health state, and let (b, y^*) be the certainty equivalent of $[(b,x),p;(b,z),\mathbf{1}-p]$ as shown in (23). I assume that $x > z$. The power QALY model (4) and Equation (35) imply that

$$k \cdot H(b) \cdot (y^*)^r = w(.5) \cdot k \cdot H(b) \cdot x^r + (\mathbf{1}-w(\mathbf{.5})) \cdot k \cdot H(b) \cdot z^r. \quad (43)$$

Therefore $$(y^*)^r = w(.5) \cdot x^r + (\mathbf{1}-w(\mathbf{.5})) \cdot z^r \quad (44)$$

and $$y^* = \left[w(.5) \cdot x^r + (1-w(.5)) \cdot z^r\right]^{1/r} \quad (45)$$

To estimate r, collect data for certainty equivalents with varying values of x and z. Let the certainty equivalents serve as the dependent variable, and the values of x and z serve as the predictor variables in a nonlinear regression that solves for estimates of $w(.5)$ and r. This solves UA Problem 2 by means of certainty equivalents under RDU assumptions. Although the estimate of $w(.5)$ may have theoretical interest, it is not directly relevant to the utility assessment. For purposes of utility assessment, the estimate of $w(.5)$ is needed only to provide an accurate estimate of the risk parameter r.

To illustrate this method, Table 2 contains hypothetical certainty equivalents data for five even-chance gambles. As shown to the right of Table 2, the fit of Equation (45) to these data yields an estimate of $r = .90$ and $w(.5) = .44$. Notice that the RDU solution for these data is less risk averse than the EU solution for these same data. This is due to the fact that the apparent risk aversion in the certainty equivalents has been absorbed into a slightly pessimistic value of $w(.5)$, i.e., $w(.5) < .5$, as well as in the slightly risk averse value of r.

To solve UA Problem 4 under RDU assumptions, we need to combine this method for estimating r with risk-adjustment of PTTOs. One must first estimate r as explained in the previous paragraph. To assess $H(b)$ for a health state b, one must determines a TTO for b as shown in (20). RDU theory and the power QALY model then imply (28) and (29), the latter equation being the risk adjustment needed to convert the PTTO to $H(b)$. One repeats this procedure for each health state b, thereby solving UA Problem 4 under RDU assumptions.

Wakker and Deneffe (1996) have recently presented a method of utility assessment that is based on quite different principles from any of the methods described in this section. Their method avoids distortions in utility that can be produced by nonlinear probability weighting, as do the methods described in this section. It also has the advantage that it does not require that one assume that all individuals have the same probability weighting function, nor does it require that the utility function be drawn from a known class of parametric functions like the power functions. I have omitted the method of Wakker and Deneffe (1996) from this chapter because the methods discussed here are essentially revisions of standard EU assessment methods, whereas Wakker and Deneffe's method is not a variant of a standard EU method. Rather, Wakker and Deneffe's method harkens back to the basic process by which equally spaced points (so-called standard sequences) are constructed in fundamental measurement theory (Krantz et al., 1971).

Summary of RDU Assessment Methods

I will summarize the assessment procedures that are sufficient to solve UA Problems 1 - 4 under RDU assumptions.

UA Problem 1. Assess the utilities of holistic outcomes.

- The standard gamble method provides the only solution when no attribute structure is specified for the outcomes. Probability equivalents p^* must be transformed to corresponding probability weights $w(p^*)$ in order to carry out the utility assessment.

UA Problem 2. Assess the utility of survival duration in some fixed health state.

- The standard gamble method yields probability equivalents for a series of survival durations. The probability equivalents must be transformed to probability weights which serve as utilities after multiplying by a scaling constant, e.g., 100. A piecewise linear approximation to the utility function can be inferred from these utilities, or a power function can be fit by means of nonlinear regression.

- Alternatively, certainty equivalents can be collected, and a power function can be fit to these equivalents by means of a nonlinear regression procedure. This approach yields estimates of $w(.5)$ and r.

UA Problem 3. Assess the values of H(b), assuming the validity of the linear QALY model (1):

- The standard gamble method yields probability equivalents p^* for each health state b. These probability equivalents must be transformed to corresponding probability weights $w(p^*) = H(b)$.
- Alternatively, proportional time tradeoffs (PTTOs) also yield values of $H(b)$ for each health state b.

UA Problem 4. Assess the values of r and H(b), assuming the validity of the power QALY model (4):

- The standard gamble method yields probability equivalents p^* for individual survival durations. The probability equivalents must be transformed to probability weights which serve as utilities after multiplying by a scaling constant. A power function can be fit to these utilities by means of nonlinear regression, thereby providing an estimate of r. To assess values of $H(b)$, one finds probability equivalents p^* for each health state b. These probability equivalents must be transformed to corresponding probability weights $w(p^*) = H(b)$.
- Alternatively, certainty equivalents can be collected, and a power function fit to these certainty equivalents as in UA Problem 2. PTTOs must be adjusted for risk posture as in Equation (29) to yield values of $H(b)$ for each health state b.

Note that all of the assessments that involve standard gambles require that one can determine $w(p^*)$ for each probability equivalent p^*. The assessments that involve certainty equivalents do not require that the probability weighting function w be known, but they assume that the utility of survival duration is a power function. One can replace this assumption by a different parametric class of functions, e.g., one can assume instead that the utility of survival duration is an exponential function.

Table 3
High and low outcomes
for 6 even chance gambles.

Stimulus Gamble	High Outcome	Low Outcome
1	12 years	0 years
2	12 years	1 year
3	12 years	4 years
4	24 years	0 years
5	24 years	2 years
6	24 years	8 years

An Empirical Example of RDU QALY Assessment[9]

To illustrate the implications of RDU utility assessment on QALY assessment, it is informative to consider some empirical results. Data from Miyamoto and Eraker (1988) will be used to estimate the parameters of the power QALY model (4) under EU and RDU assumptions. We will examine the parameter estimates to see how EU and RDU theory differ in their interpretation of the same data.

Miyamoto and Eraker (1988) reported an experiment in which a sample of medical patients judged certainty equivalents of even-chance gambles for survival duration. Table 3 lists the high and low outcomes for six even-chance gambles that were presented as stimuli in the “power” condition of the experiment. Subjects in this condition were asked to judge the certainty equivalents of these gambles assuming full health, and a second time, assuming life with their current symptoms. Each subject also made 8 TTO judgments as in Equation (20). The time tradeoffs were judged with respect to 15, 16, 20, and 24 years with current symptoms.

9 The work in this section is part of a larger project on non-EU utility assessment that is being developed jointly with Richard Gonzalez and Jon Treadwell.

Table 4
Certainty Equivalents Data for 4 Subjects

		Subject 1		Subject 2		Subject 7		Subject 23	
Health	Stimulus	Trial		Trial		Trial		Trial	
State	Gamble	1	2	1	2	1	2	1	2
Full	1	6.00	6.00	6.00	5.50	10.00	10.00	6.50	7.00
Full	2	7.00	6.00	7.00	7.00	9.50	10.50	4.50	8.50
Full	3	9.00	8.00	9.00	9.00	10.50	10.75	7.50	8.50
Full	4	11.00	12.00	10.00	12.00	20.50	21.00	5.50	9.50
Full	5	11.00	12.00	13.00	13.00	20.50	21.50	7.50	12.50
Full	6	11.00	12.00	17.00	18.00	21.00	21.00	13.50	13.50
Current	1	6.00	7.00	5.50	6.00	10.00	10.50	6.50	6.00
Current	2	7.00	6.00	5.50	6.50	10.50	10.50	6.50	6.50
Current	3	8.00	8.00	9.50	9.00	11.00	10.75	7.50	8.50
Current	4	12.00	12.00	7.50	12.00	21.00	21.00	8.50	12.50
Current	5	12.00	9.00	11.00	13.00	21.50	21.50	9.50	9.50
Current	6	16.00	16.00	14.00	17.00	20.00	22.00	13.50	15.50

Every certainty equivalence judgment and TTO judgment was replicated twice. A more complete description of the subjects and experimental procedure is given in Miyamoto and Eraker (1988).

Table 4 displays the certainty equivalence data for Subjects 1, 2, 7, and 23. The gambles that elicited these certainty equivalents are the six gambles shown in Table 3. Subjects were instructed to assume that survival would be accompanied by full health for the gambles in the upper half of Table 4, and by current symptoms for the gambles in the lower half of the table. The complete data is given for these subjects in order that the interested reader can reproduce the utility analysis as it is presented here. For the sake of brevity, the TTO data are not presented, but the mean PTTOs are listed in Table 5, column 5.

Let $\hat{r}_{EU}$ stand for an estimate of the risk parameter of the power QALY model (4) computed under the assumption that EU theory is valid. To determine $\hat{r}_{EU}$ for each subject, let Equation (27) define the model to be estimated in a nonlinear regression. The certainty equivalents in Table 4 serve as the dependent variable, and the high and low outcomes in Table 3 serve as the predictor variables. The starting value of $r = 1$ was employed in a nonlinear, least squares regression. Column 2 of Table 5 displays the estimates of r calculated under EU assumptions.

Let $\hat{r}_{RDU}$ stand for an estimate of r that is computed under the assumption that RDU is valid. To determine $\hat{r}_{RDU}$ for each subject, let Equation (45) define the model to be estimated in a nonlinear regression.

The certainty equivalents in Table 4 serve as the dependent variable, and the high and low outcomes in Table 3 serve as the predictor variables. The risk parameter *r* and the probability weight *w(.5)* are the parameters to be estimated in the regression. The starting values of $r = 1$ and $w(.5) = .5$ were employed in a nonlinear, least squares regression. Columns 3 and 4 of Table 5 display the estimates of *r* and *w(.5)* calculated under RDU assumptions.

Let a = *full health* and b = *current symptoms*. Let $H(a) = 1$. Assuming the linear QALY model (1), the PTTO equals *H(b)* as shown in Equation (22). Column 5 of Table 5 shows the mean PTTO for each subject averaged over 8 PTTO judgments. Note that the means shown in Column 5 of Table 5 are estimates of *H(b)* for the linear QALY model under either EU or RDU assumptions because Equation (22) is implied by the linear QALY model under either EU or RDU assumptions.

Suppose, now, that we drop the assumption of linearity and assume instead that a power QALY model (4) holds. EU theory and RDU theory both imply that the calculation of *H(b)* must adjust the PTTO for the risk parameter *r* as shown in Equation (29), but the RA-PTTOs will differ because EU and RDU theory yield different estimates of *r*. Columns 6 and 7 of Table 5 show the RA-PTTOs for EU and RDU theory, respectively[10].

In other words, Column 6 shows $\left[\text{mean}(y^*/x)\right]^{\hat{r}_{EU}}$, and Column 7 shows $\left[\text{mean}(y^*/x)\right]^{\hat{r}_{RDU}}$.

Let us compare the utility assessment under EU and RDU assumptions for these four subjects. For Subject 1, the utility of survival duration is slightly risk averse under EU assumptions ($\hat{r}_{EU}$ = .89). The RDU analysis, however, suggests that this subject is actually rather risk seeking ($\hat{r}_{RDU}$ = 1.60). The reason for this discrepancy is that the RDU analysis finds that this subject is rather pessimistic with respect to the .5 probability ($w(.5) = .31$). The preference behavior that the EU analysis attributes to risk

Table 5

	EU Theory	*RDU Theory*		*Linear QALY*	*EU Theory*	*RDU Theory*
Subject	$\hat{r}_{EU}$	$\hat{r}_{RDU}$	*w(.5)*	*PTTO*	*RA PTTO*	*RA PTTO*
1	.89	1.60	.31	.56	.60	.40
2	.89	.65	.58	.76	.78	.84
7	5.15	2.61	.70	.89	.54	.73
23	.72	1.06	.38	.55	.65	.53

aversion is associated so strongly with pessimism in the RDU analysis that the curvature in the utility function is reversed from risk aversion to risk seeking. These differing assessments of risk posture impact the assessment of *H(b)*, the utility of current symptoms. Assuming risk neutrality, $H(b) = PTTO = .56$. Risk adjustment with respect to $\hat{r}_{EU}$ yields $H(b) = .60$. Risk adjustment with respect to $\hat{r}_{RDU}$ yields $H(b) = .40$, a considerably lower utility for *b*. Although the lower value of *H(b)* found under RDU assumptions is a direct result of the fact that $\hat{r}_{RDU}$ is rather risk seeking, (see Figure 5), we should note that the more basic reason for the discrepancy between the EU and RDU assessment of *H(b)* is that the latter theory allows for pessimism in the probability weighting. Nonlinear probability weighting impacts both the assessment of risk posture and the assessment of health state utilities by TTO methods.

Under EU assumptions, Subject 2 is precisely as risk averse as Subject 1 ($\hat{r}_{EU} = .89$), but the RDU analysis found Subject 2 to be slightly optimistic with respect to the .5 probability ($w(.5) = .58$). Therefore, unlike Subject 1, Subject 2 is more risk averse under the RDU analysis than under the EU analysis. Consequently, *H(b)* assessed under RDU assumptions is greater than the estimate of *H(b)* found under either the linear QALY model or the power QALY model and EU theory.

---

10 Technically, it would be better to apply the power transformations to individual PTTOs prior to averaging them, and then average these transformed PTTOs, but the risk-adjusted mean PTTOs reported in Table 5 are within ± .01 of the mean RA-PTTOs that would be computed by this alternative method.

Subject 7 is an example of an individual who appears to be extremely risk seeking under EU assumptions ($\hat{r}_{EU}$ = 5.15). When nonlinear probability weighting is taken into account in the RDU analysis, the subject is considerably less risk seeking, although still risk seeking ($\hat{r}_{RDU}$ = 2.61), and is found to be rather optimistic (*w(.5) = .70*). Whereas the EU analysis yields a RA-PTTO of .54 which is far below the PTTO of .89 found under linear QALY assumptions, the RDU analysis yields a RA-PTTO of .73 which is approximately midway between .54 and .89. Absorbing some of the apparently risk seeking preferences into an optimistic estimate of *w(.5)* reduces the discrepancy between the estimate of *H(b)* found under linear QALY model and the estimate of *H(b)* found under EU assumptions and the power QALY model.

Finally, Subject 23 appears to be somewhat risk averse under EU assumptions ($\hat{r}_{EU}$ = .72), but the RDU analysis attributes the apparently risk averse preferences almost entirely to pessimism ($\hat{r}_{RDU}$ *= 1.06, w(.5) = .38*). Therefore the utility of health state *b* is almost identical under the linear QALY model (*H(b) = .55*) and under the power QALY model and RDU assumptions (*H(b) = .53*), whereas it is somewhat higher under the power QALY model and EU asssumptions (*H(b) = .65*).

These examples illustrate several differences between utility assessments under EU and RDU assumptions. First, in general $\hat{r}_{EU} < \hat{r}_{RDU}$ when *w(.5) < .5*, and conversely, $\hat{r}_{EU} > \hat{r}_{RDU}$ when *w(.5) > .5*. Richard Gonzalez, Jon Treadwell, and I have examined certainty equivalents data for 65 utility functions for survival duration, and have found only one exception to this pattern. Equation (29) shows that the PTTO must be adjusted for risk posture in order to estimate *H(b)*. As we have seen here, nonlinear probability weighting has a systematic impact on the degree of risk aversion or risk seeking that will be found in the assessed utility function. Putting these two findings together, we can see that if 1 > (y*/x) > 0, then in general,

$$\left(\frac{y^*}{x}\right)^{\hat{r}_{EU}} > \left(\frac{y^*}{x}\right)^{\hat{r}_{RDU}} \text{ if and only if } \hat{r}_{EU} < \hat{r}_{RDU}$$

$$\text{if and only if } \quad w(.5) < .5.$$

Because $(y^*/x)^{\hat{r}_{EU}}$ is the EU assessment of *H(b)* and $(y^*/x)^{\hat{r}_{RDU}}$ is the RDU assessment of *H(b)*, this shows that the EU assessment of *H(b)* will exceed the RDU assessment of *H(b)* for individuals who are pessimistic with respect to the .5 probability.

What these examples show is that nonlinear probability weighting impacts both the measure of risk posture, *r*, and the RA-TTO. Although this example is confined to estimation of risk posture from certainty equivalents data, I believe that analogous relationships will be found in standard gambles data. In other words, if standard gambles data are collected that allow one to estimate the form of nonlinear probability weighting for individual subjects, the RDU utility assessment will be relatively more risk seeking than the EU utility assessment for individuals who are pessimistic with respect to the probabilities in the utility assessment. For health state utilities, the RDU estimates of *H(b)* should be lower than the EU estimates of *H(b)* for individuals who are pessimistic with respect to the probabilities in the utility assessment. In effect, nonlinear probability weighting should impact measures of risk posture and health state utility regardless of whether utilities are assessed by standard gambles or certainty equivalents and time tradeoffs.

Conclusions

The assessment of utilities under EU assumptions is strongly influenced by the assumption that probabilities contribute linearly to the utility of a lottery. This is certainly true of the standard gamble where the probability equivalent of an outcome is interpreted as the utility of that outcome (after possible rescaling--see Equation (11)). In the method of certainty equivalents, the .5 probability that appears in the stimulus gamble is assumed to carry a .5 weight in the evaluation of the utility of an outcome (see Equation (27)). The only exception is the assessment of health state utilities by means of TTOs under the assumption that the utility of survival duration is linear. If the linear QALY model is assumed to be descriptively valid, then the PTTO equals the health state utility under EU or RDU assumptions. In effect, if one makes the strong assumption that the utility of survival duration is linear, one can use TTOs to assess health state utilities and thereby avoid the assumption that the perception of probability is linear. It can be shown, however, that the utility of survival duration is typically not linear (Miyamoto & Eraker, 1985, 1989), and therefore, descriptive accuracy requires risk-adjustment of the PTTOs. Risk-adjusted TTOs are influenced by nonlinear probability perception because nonlinear probability weighting affects the estimation of the risk parameter *r* (see Equations (29)).

RDU theory provides a theoretical basis for taking nonlinear probability weighting into account in the process of utility assessment. Given a probability weighting function *w*, the standard gamble method can be applied under RDU assumptions to assess the utilities of individual outcomes, health state utilities, or the utility of continua like survival duration. At present, only population-based estimates of *w* are available, and it can be questioned whether a single weighting function is sufficiently close to the weighting functions of individuals. It is likely that research in the near future will determine whether it is possible to estimate weighting functions for individuals from the kinds of small data sets that are common in health utility analysis. This chapter also describes an approach to QALY measurement based on certainty equivalents and TTOs. Richard Gonzalez, Jon Treadwell, and I are currently investigating this approach to QALY measurement. An advantage of this approach is that it only requires the estimation of one probability weight, *w(.5)*, but it requires that one assume that the utility of survival duration is drawn from a specific class of parametric utility functions, like power functions or exponential functions.

As was shown in the examples of utility assessment, nonlinear probability weighting affects both the measurement of risk posture and the assessment of health state utilities. Pessimistic probability weights reduce the degree of risk aversion that is found in the assessments of the utility of survival duration. Not only do changes in risk aversion affect utility tradeoffs between short and long-term survival, they also affect assessments of health state utility through the process of risk adjustment of TTOs. One goal of future research should be to investigate whether decision analyses are sensitive to descrepancies between EU and RDU assessments of health state utilities.

Another goal of future research will be the further generalization of utility assessment methods to cumulative prospect theory (CPT) of Tversky and Kahneman (1992). The principal difference between CPT and RDU theory is that CPT postulates that outcomes are perceived as gains or losses relative to a reference level rather than as absolute levels of wealth or health. Furthermore, CPT postulates that the utility representation for lotteries can differ depending on whether the outcomes of a lottery are exclusively nonlosses (gains or "zero" outcomes), exclusively nongains (losses or "zero" outcomes), or a mixture of gains and losses. The CPT representation for lotteries that are nonlosses is isomorphic to the RDU representation, as is the CPT representation for lotteries that are nongains. Hence the methods for RDU assessment that were described in this chapter remain valid for CPT when the domain consists exclusively of nonloss outcomes or of nongain outcomes. The main issues that arise in the generalization of the present work to CPT are, first, the development of a methodology for identifying reference levels in health domains, and second, the development of assessment methods that take into account preferences for lotteries whose outcomes are mixtures of losses and gains.

References

Abdellaoui, M. (unpublished manuscript). Parameter-free elicitation of utilities and probability weighting functions.

Bayoumi. A. M., & Redelmeier, D. A. (1996). Correcting distorted utilities: An application of cummulative prospect theory to medical economics [abstract]. *Medical Decision Making*, 16, 457.

Bleichrodt, H., & Pinto, J. L. (unpublished manuscript). A parameter-free elicitation of the probability weighting function in medical decision making.

Bleichrodt, H., Wakker, P., & Johannesson, M. (1997). Characterizing QALYs by risk neutrality. *Journal of Risk and Uncertainty*, 15, 107-114.

Bostic, R., Herrnstein, R. J., & Luce, R. D. (1990). The effect on the preference reversal phenomenon of using choice indifferences. *Journal of Economic Behavior and Organization*, 13, 193-212.

Camerer, C. F. (1989). An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty*, 2, 61-104.

Camerer, C., & Ho. T. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty*, 8, 167-196.

Cher, D. J., Miyamoto, J. M., & Lenert, L. A. (1997). Risk adjustment of markov process models: Importance in individual decision making. *Medical Decision Making*, 17, 340-350.

Drummond, MF, O'Brien, B, Stoddart, GL, & Torrance, GW (1997). *Methods for the economic evaluation of health card programmes*, (second edition). New York: Oxford University Press.

Edwards, W. (Ed.) (1992). *Utility theories: Measurements and applications*. Boston: Kluwer.

Fanshel, S., & Bush, J. W. (1970). A health-status index and its application to health-services outcomes. *Operations Research*, 18, 1021-1066.

Fischer, G. W., & Hawkins, S. A. (1983). Strategy compatibility, scale compatibility, and the prominence effect. *Journal of Experimental Psychology: Human Perception and Performance*, 19, 580-597.

Fishburn, P. C. (1982). *The foundations of expected utility theory*. Dordrecht, NL: D. Reidel.

Fishburn, P. C. (1989). Retrospective on the utility theory of von Neumann and Morgenstern. *Journal of Risk and Uncertainty*, 2, 127-157.

Froberg, D. G., & Kane, R. L. (1989a). Methodology for measuring health-state preferences--I: Measurement strategies. *Journal of Clinical Epidemiology*, 42(4), 345-354.

Froberg, D. G., & Kane, R. L. (1989b). Methodology for measuring health-state preferences--II: Scaling methods. *Journal of Clinical Epidemiology*, 42(5), 459-471.

Froberg, D. G., & Kane, R. L. (1989c). Methodology for measuring health-state preferences--III: Popluation and context effects. *Journal of Clinical Epidemiology*, 42(6), 585-592.

Froberg, D. G., & Kane, R. L. (1989d). Methodology for measuring health-state preferences--IV: Progress and a research agenda. *Journal of Clinical Epidemiology*, 42(7), 675-685.

Fryback, D. G. (in press). The QALY model: Utilities for cost-utility analysis in health care. In J. Shanteau, B. Mellers, & D. Schum (Eds.), *Decision research from Bayesian approaches to normative systems: Reflections on the contributions of Ward Edwards*. Norwell, MA: Kluwer Academic Publishers.

Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein, M. C. (1996). *Cost-effectiveness in health and medicine*. New York: Oxford University Press.

Gonzalez, R. (1993). Estimating the weighting function. Paper presented at the 26th Annual Mathematical Psychology Meeting, Oklahoma City, Oklahoma.

Gonzalez, R. & Wu, G. (unpublished manuscript). On the form of the probability weighting function.

Karni, E., & Safra, Z. (1990). Rank dependent probabilities. *Economics Journal*, 100, 187-195.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 276-287.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*, 341-350.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives.* New York: Wiley.

Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement*, (Vol. 1). New York: Academic Press.

Llewellyn-Thomas, H., Sutherland, H. J., Tibshirani, R., Ciampi, A., Till, J. E., & Boyd, N. F. (1982). The measurement of patients' values in medicine. Medical Decision Making, 2, 449-462.

Luce, R. D. (1992). Where does subjective expected utility fail descriptively? *Journal of Risk and Uncertainty*, 5, 5-27.

McNeil, B. J., & Pauker, S. G. (1982). Optimizing patient and societal decision making by the incorporation of individual values. In R. L. Kane & R. A. Kane, (Eds.), *Values and long-term care* (pp. 215-230). Lexington, MA: D. C. Heath.

McNeil, B. J., Weichselbaum, R., Pauker, S. G. (1978). Fallacy of the five-year survival in lung cancer. *New England Journal of Medicine*, 299, 1397-1401.

Miyamoto J. M. (1992). Generic analysis of utility models. In W. Edwards (Ed.), *Utility: Theories, measurement, and applications*. Boston: Kluwer, 73-106.

Miyamoto, J. M. (in press) Quality-adjusted life years (QALY) utility models under expected utility and rank-dependent utility assumptions. *Journal of Mathematical Psychology*.

Miyamoto, J. M., & Eraker, S. A. (1985). Parameter estimates for a QALY utility model. *Medical Decision Making*, *5*, 191-213.

Miyamoto, J. M., & Eraker, S. A. (1988). A multiplicative model of the utility of survival duration and health quality. *Journal of Experimental Psychology: General*, *117*, 3-20.

Miyamoto, J. M., & Eraker, S. A. (1989). Parametric models of the utility of survival duration: Tests of axioms in a generic utility framework. *Organizational Behavior and Human Decision Processes*, *44*, 166-202.

Miyamoto, J. M., Wakker, P., Bleichrodt, H., & Peters, H. (in press). The zero condition: A simplifying assumption in QALY measurement. *Management Science*.

Patrick, D. L., Starks, H. E., Cain, K. C., Uhlmann, R. F., & Pearlman, R. A. (1994). Measuring preferences for health states worse than death. *Medical Decision Making*, 14, 9-18.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Oranization*, 3, 323-343.

Quiggin, J. (1993). *Generalized expected utility theory--The rank-dependent model*. Boston: Kluwer.

Quiggin, J. & Wakker, P. (1994). The axiomatic basis of anticipated utility theory: A clarification. *Journal of Economic Theory*, 64, 486-499.

Roberts, F. S. (1979). *Measurement theory with applications to decisionmaking, utility, and the social sciences.* London: Addison-Wesley. Pages 13 - 47, 101 - 148.

S-Plus for Windows Versions 3.3. Distributed by MathSoft, Inc., Data Analysis Products Division, 1700 Westlake Avenue N, #500, Seattle, WA 98109-9891.

Slovic, P., and Lichtenstein, S. (1968). Relative importance of probabilities and payoffs in risk taking. *Journal of Experimental Psychology*, 78, 1-18.

Slovic, P., Lichtenstein, S., and Fischhoff, B. (1988). Decision Making. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.). *Steven's Handbook of Experimental Psychology*, (Vol. 2). 673-738.

Sox, H. C., Blatt, M. A., Higgins, M. C., & Marton, K. I. (1988). *Medical decision making*. Boston: Butterworths.

SPSS for Windows, Version 7.5, distributed by SPSS Inc., 444 N. Michigan Avenue, Chicago, Illinois, 60611.

Stalmeier, P. F. M., Bezembinder, T. G. G., and Unic, I. J. (1996). Proportional heuristics in time tradeoff and conjoint measurement. *Medical Decision Making*, 16, 36-44.

Sutherland, H. J., Llewelynn-Thomas, R. N. J., Boyd, N. F., & Till, J. E. (1982). Attitudes toward quality of life: The concept of "Maximal Endurable Time." *Medical Decision Making*, 2, 299-309.

Torrance, G. W. (1986). Measurement of health state utilities for economic appraisal: A review. *Journal of Health Economics*, 5, 1-30.

Torrance, G. W., Thomas, W. H., & Sackett, D. L. (1972). A utility maximization model for evaluation of health care programs. *Health Services Research, 7,* 118-133.

Tversky A, & Kahneman D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297-323.

Tversky, A., Sattath, S., & Slovic, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, 95, 371-384.

von Neumann, J., & Morgenstern, O. (1944). *The theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral decision theory*. New York: Cambridge University Press.

Wakker, P. P. (1989). *Additive representations of preferences : A new foundation of decision analysis.* Boston: Kluwer.

Wakker, P., & Deneffe, D. (1996). Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management Science*, 42, 1131-1150.

Wakker P, & Stiggelbout A. (1995). Explaining distortions in utility elicitation through the rank-dependent model for risky choices. *Medical Decision Making*, 15, 180-186.

Weinstein, M. C., Fineberg, H. V., Elstein, A. S., Frazier, H. S., Neuhauser, D., Neutra, R. R., & McNeil, B. J. (1980). *Clinical decision analysis*. Philadelphia, PA: W. B. Saunders.

Wu, G., & Gonzalez, R. (1996). Curvature of the probability weighting function. *Management Science*, 42, 1676 - 1690.