KM34(1990)
pag 53 - 64

TESTING THE RANK-ORDER APPROACH TO UTILITY FOR MONEY AND HEALTH STATES

by Eugene Canjels* and Hans Peters†

Abstract

Results are reported of a preliminary experiment testing the rank-order approach to utility for money and health states. In particular, it is tested not only whether the rank-order approach is valid, but also to which extent probability distortion associated with risky outcomes, depends on the involved prizes. The rank-order approach is an alternative to the Von Neumann-Morgenstern theory and allows a distinction between utilities for risky and riskless events. The adoption of this approach may have consequences in particular for the use of the standard gamble method for measuring utility indices of health states. These, in turn, are of practical relevance for the determination of so-called quality-adjusted life years used to evaluate health care programs. Our results are not yet conclusive, but will be useful in designing future experiments, and in particular indicate that in certain ranges the standard gamble method should be used with great care, especially so if 'bad' prizes are very likely.

*Department of Economics, University of Limburg, P.O. Box 616, 6200 MD Maastricht, The Netherlands, tel.: 043-888821

†Faculty of General Sciences, Department of Mathematics, University of Limburg, P.O. Box 616, 6200 MD Maastricht, The Netherlands, tel.: 043-887503

1 Introduction

Since its introduction by Von Neumann and Morgenstern (1947), expected utility theory has been widely accepted and applied in economics and other social sciences. The theory is mathematically simple and elegant, and can be based on a number of appealing axioms (see also Herstein and Milnor (1953)). Empirical investigations, however, have resulted in many violations of the theory. Well-known is the so-called Allais paradox (see, e.g., Allais (1979)): this is an example of a risky choice situation where subjects tend to violate expected utility maximization.

Scientists in this area, theorists as well as empiricists, can be roughly divided into two groups. First, there are those who think that individuals in their right minds would choose such as to maximize their expected utility, and that observed violations of the theory are caused by biases resulting from the experiment, test, or observations. In other words, they believe that, provided certain test conditions are fulfilled, expected utility theory does have descriptive value. According to scientists in this group, one should spend one's effort in trying to design experiments that rule out these biases as much as possible. Second, alternative theories have been developed by researchers who think that expected utility theory does not have adequate descriptive power. We mention, in particular, regret theory (Loomes and Sugden (1982)), prospect theory (Kahneman and Tversky (1979)), and the so-called rank-order approach (Quiggin (1982), Yaari (1987)). A recent survey on decision making under uncertainty is Machina (1987). Jaffray (1986) summarizes some important experimental findings and difficulties in this area.

The validity of expected utility theory as a descriptive theory is of practical importance for health care planning, since the commonly used methods for assessing individuals' preferences with respect to health states are based on the Von Neumann-Morgenstern theory. The resulting utility indices are needed to determine so-called qaly's (quality adjusted life years), and these, in turn, play an important role in deciding how to distribute financial resources over various health care programs. See, e.g., Pliskin et al. (1980), and Torrance (1986).

These commonly used methods are: Visual Analogue (VA), Standard Gamble (SG), and Time Trade-off (TT). Let us fix an individual's utility for being dead at 0 and for being in complete health at 1. We want to assess that individual's utility u for a certain state of health H. In the VA method, that individual would be asked, for instance, to draw a line of length x between 0 and 1 as an indication of u. In the SG method, the individual would be asked to mention a number p between 0 and 1, such that (s)he would be indifferent between undergoing a surgical operation resulting with probabilities p and $1-p$ in complete health and death, respectively, and remaining in state H for the rest of his/her life. In the TT method, finally, the individual would have to name a number t between 0 and (say) 10, such that (s)he would be indifferent between 10 years of remaining lifetime in state H and t years in complete health. Under expected utility theory, an individual answers in a consistent manner if and only if $x = p = t/10$. (Here, it is assumed that - ceteris paribus -utility is linear with length of life.) See Torrance (1986) or Loomes (1988) for a more extensive discussion of these methods. Both authors survey a number of experiments where these three methods were used, showing systematic discrepancies between the generated utility indices: that is, x, p and $t/10$ are seldomly equal. We mention Torrance (1976), Bombardier et al. (1982), Llewllyn-Thomas et al. (1984), Torrance et al. (1982).

Several explanations have been offered to account for the discrepancies between Visual Analogue, Standard Gamble, and Time Trade-off experimentation results. Here, we focus on an argument that distinguishes between VA and TT on the one hand, and SG on the other hand: in the Standard Gamble method, subjects are asked to decide under risk, whereas there is no risk involved in the other two methods. Bombardier et al. (1982, p. 152) found that "almost universally, the standard gamble values were higher". As a possible explanation, they suggest a

> general aversion to gambling with one's health, a 'gambling aversion' which must be distinguished from the 'risk aversion' familiar to students of decision analysis.

The authors refer to the 'risk aversion' according to the Von Neumann-Morgenstern theory, which is equivalent to concavity of the utility function u. Indeed, this cannot explain the higher level of the standard gamble values. Von Neumann-Morgenstern utility theory offers no adequate way to distinguish between decision making under certainty and under risk: it does not capture psychological phenomena like the 'gambling aversion' above. A related criticism is the unavoidable twinship in Von Neumann-Morgenstern utility theory, of risk aversion and decreasing marginal utility (cf. Peters (1988)): both are equivalent to concavity of the utility function.

The mentioned alternative theories might offer solutions to this problem. Loomes (1988) describes how regret theory can be used as a theoretical framework for measuring individuals' preferences with respect to health conditions. Here, we focus on what is usually called the rank-order approach. This theory offers a way to distinguish between decision making under certainty and under risk; it contains expected utility as a special case. It tries to capture our intuition that people tend to subjectively *distort* probabilities depending on the *prizes* that are associated with these probabilities. We report the results of a pilot experiment in which these *prizes* were either money or health states. The aim of this experiment was threefold. First, to test the validity of the rank-order approach. Second, to compare results when prizes are monetary with results for health conditions as prizes. Third, to reach a conclusion on the usefulness of, in particular, the standard gamble method for the determination of utility indices and, ultimately, *qaly's* in connection with health states and health care programs.

The organization of the remainder of the paper is as follows. Section 2 describes the theoretical framework. Section 3 describes the theoretical and practical design of the experiment. Section 4 gives the results and conclusions.

2 The Rank-Order Approach

We will describe the rank-order approach to utility in a form that is convenient for our experimental set-up. We follow Lefoll et al. (1988) who were also our main inspiration in designing the experiment.

Let $x_1, x_2, \ldots, x_n$ be a number of alternatives. The x_i may be monetary prizes but also other things like health states. We suppose an individual has a utility function u such that

$$u(x_1) \leq u(x_2) \leq \ldots \leq u(x_n). \tag{1}$$

E.g., if x_i is money, then (for an average individual) $x_1 \leq \ldots \leq x_n$. The cardinal utility function u is meant to describe the individual's preferences in case of certainty. Now let $p_1, p_2, \ldots, p_n$ be nonnegative real numbers with $\sum_{i=1}^{n} p_i = 1$, and let $\varphi : [0,1] \to [0,1]$ be a monotonically nondecreasing function with $\varphi(0) = 0$, $\varphi(1) = 1$. Let L denote the lottery in which the individual receives x_i with probability p_i. We define

$$\begin{aligned} U_\varphi(L) &= u(x_1) + \varphi(p_2 + p_3 + \ldots + p_n)(u(x_2) - u(x_1)) \\ &\quad + \varphi(p_3 + \ldots + p_n)(u(x_3) - u(x_2)) + \ldots \\ &\quad + \varphi(p_n)(u(x_n) - u(x_{n-1})). \end{aligned} \tag{2}$$

According to the ***rank-order approach***, $U_\varphi(L)$ is the individual's utility for the lottery L, given that (s)he 'distorts' probabilities in a way described by φ: the function φ represents the individual's attitude towards risk, and will be called *distortion function*.

It is important to realize that for formula (2) to make sense, prizes must be ordered increasingly, as in formula (1). Formula (2) can then be read as follows: the individual enjoys $u(x_1)$ with certainty - in view of (1); he additionally enjoys $u(x_2) - u(x_1)$ with probability $\varphi(\sum_{i=2}^n p_i)$; he additionally enjoys $u(x_3) - u(x_2)$ with probability $\varphi(\sum_{i=3}^n p_i)$, etc. If φ is identity, we have

$$U_{id}(L) = p_1u(x_1) + p_2u(x_2) + \ldots + p_nu(x_n)$$

which is the expected (Von Neumann-Morgenstern) utility of the lottery L.
The rank-order approach is an improvement of the prospect theory of Kahneman and Tversky (1979), which would put $U_\varphi(L) = \sum_{i=1}^n \varphi(p_i)u(x_i)$, violating first-order stochastic dominance: this is avoided by the rank-order approach. For an informal treatment, see Wakker (1988). The rank-order approach is backed up by axiomatizations: see Schmeidler (1984), Yaari (1987), Gilboa (1987), and Wakker (1989).

While in the Von Neumann-Morgenstern utility theory risk aversion is equivalent to *concavity* of the utility function u, in the rank-order approach risk aversion is usually characterized as *convexity* of the distortion function φ, see for instance Chew et al. (1987). From now on, unless stated otherwise, risk aversion will always be risk aversion according to the rank-order approach. Intuitively, a risk averse individual tends to distort downwards the probabilities of receiving particularly good prizes. Note that, theoretically, a risk averse individual does not necessarily exhibit decreasing marginal utility for sure outcomes: the latter is equivalent to concavity of the utility function u, and is in principle independent of properties of the distortion function.

3 The Experimental Design

During the experiment, each subject had to give 18 valuations. Half of these were concerned with money, the other half with health.

The monetary valuations are numbers m_i ($i = 1, 2, \ldots, 9$) between 0 and 1. The utility of obtaining 0 Dutch guilders was fixed at 0, the utility of receiving 1000 Dutch guilders was fixed at 1. The meaning of the numbers m_i was as follows:

m_1 = utility of receiving Dfl. 300
m_2 = utility of receiving Dfl. 500
m_3 = utility of receiving Dfl. 700.

Let $[x; p,\ y; 1-p]$ denote the lottery in which the subject receives x with probability p and y with probability $1-p$.

m_4 = utility of the lottery [Dfl. 0; 0.3, Dfl. 500; 0.7]
m_5 = utility of the lottery [Dfl. 0; 0.5, Dfl. 500; 0.5]
m_6 = utility of the lottery [Dfl. 0; 0.7, Dfl. 500; 0.3]
m_7 = utility of the lottery [Dfl. 500; 0.3, Dfl. 1000; 0.7]
m_8 = utility of the lottery [Dfl. 500; 0.5, Dfl. 1000; 0.5]
m_9 = utility of the lottery [Dfl. 500; 0.7, Dfl. 1000; 0.3].

The valuations of health conditions are numbers h_i ($i = 1, 2, \ldots, 9$) also between 0 and 1. The utility of being dead (D) was fixed at 0, while complete health (H) was worth 1. Three health, or rather, illness conditions were described, which we name here H_1, H_2, and H_3. H_1 was clearly the worst condition, and H_3 the least bad. In these valuations, lotteries were surgical operations. The meaning of the numbers h_i was as follows:

h_1 = utility of being in health state H_1
h_2 = utility of being in health state H_2
h_3 = utility of being in health state H_3.

In each of the following lotteries, the subject had to imagine him/herself being in health state H_2.

h_4 = utility of the lottery [D; 0.3, H; 0.7]
h_5 = utility of the lottery [D; 0.5, H; 0.5]
h_6 = utility of the lottery [D; 0.7, H; 0.3]
h_7 = utility of the lottery [H_1; 0.3, H_3; 0.7]
h_8 = utility of the lottery [H_1; 0.5, H_3; 0.5]
h_9 = utility of the lottery [H_1; 0.7, H_3; 0.3].

For each subject, we postulated the existence of utility functions u^m and u^h ('m' for money, 'h' for health) and 'probability functions' φ^m and φ^h such that (2) is satisfied for the pair u^m, φ^m as well as for the pair u^h, φ^h. Per individual, three values of u^m and three values of u^h are measured, namely m_1, m_2, m_3, and h_1, h_2, h_3, respectively. With the aid of these, and by considering the individual valuations of the lotteries, for each of the distorted probabilities $\varphi^m(0.7), \varphi^m(0.5), \varphi^m(0.3), \varphi^h(0.7), \varphi^h(0.5), \varphi^h(0.3)$ two values can be calculated, to be used in the different tests.
In particular, we derive:

$u^m(0) = 0$, $u^m(300) = m_1$, $u^m(500) = m_2$, $u^m(700) = m_3$, $u^m(1000) = 1$, $m_4 = \varphi^m(0.7)m_2 \Rightarrow \varphi^m(0.7) = m_4/m_2$, similarly:
$\varphi^m(0.7) = (m_7 - m_2)/(1 - m_2)$, $\varphi^m(0.5) = m_5/m_2$, $\varphi^m(0.5) = (m_8 - m_2)/(1 - m_2)$, $\varphi^m(0.3) = m_6/m_2$, $\varphi^m(0.3) = (m_9 - m_2)/(1 - m_2)$.

Similarly for the health part of the test:

$u^h(D) = 0$, $u^h(H_1) = h_1$, $u^h(H_2) = h_2$, $u^h(H_3) = h_3$, $u^h(H) = 1$, $\varphi^h(0.7) = h_4$, $\varphi^h(0.7) = (h_7 - h_1)/(h_3 - h_1)$, $\varphi^h(0.5) = h_5$, $\varphi^h(0.5) = (h_8 - h_1)/(h_3 - h_1)$, $\varphi^h(0.3) = h_6$, $\varphi^h(0.3) = (h_9 - h_1)/(h_3 - h_1)$.

(Note that the numbers m_1 and m_3 play no role in the determination of the values of φ^m. They were included as reference points for the subjects.)

The values of φ^m and φ^h thus obtained, can be used for a number of tests. The test method used was a nonparametric procedure fit for these data, namely the Wilcoxon signed rank test. We list the hypotheses tested:

(a) Test of the null hypothesis: Consistency of probability distortion, and thus, validity of the rank-order approach, concerning the evaluation of monetary prizes.

Formally, we consider the differences (one per subject) $\Delta_1^m = m_4/m_2 - (m_7 - m_2)/(1 - m_2)$, $\Delta_2^m = m_5/m_2 - (m_8 - m_2)/(1 - m_2)$, $\Delta_3^m = m_6/m_2 - (m_9 - m_2)/(1 - m_2)$, that is, the differences between the pairs of values measured for $\varphi^m(0.7)$, $\varphi^m(0.5)$, and $\varphi^m(0.3)$, respectively. Null hypothesis: $\Delta_i^m = 0$ $(i = 1, 2, 3)$, alternative: $\Delta_i^m \neq 0$. The null hypothesis is in line with the rank-order approach.

(b) Test of the null hypothesis: Consistency of probability distortion, and thus, validity of the rank-order approach, concerning the evaluation of health states.

Formally, we consider the differences $\Delta_1^h = h_4 - (h_7 - h_1)/(h_3 - h_1)$, $\Delta_2^h = h_5 - (h_8 - h_1)/(h_3 - h_1)$, $\Delta_3^h = h_6 - (h_9 - h_1)/(h_3 - h_1)$. Null hypothesis: $\Delta_i^h = 0$ $(i = 1, 2, 3)$, alternative: $\Delta_i^h \neq 0$. The null hypothesis is in line with the rank-order approach.

(c) Tests on risk aversion: for each of the twelve lotteries involved in the experiment, we test the null hypothesis 'no risk aversion' against the alternative 'risk aversion'. Notice — see section 2 — that the absence of risk aversion, or, more precisely, risk neutrality according to the rank-order approach, 'happens' to be in line with the Von Neumann-Morgenstern utility theory, since both correspond to the distortion function being linear.

Formally, we test the null hypothesis $\Delta = \varphi(p) - p = 0$ against $\Delta = \varphi(p) - p < 0$ (which indicates risk aversion according to the rank-order approach). This test is performed for the following cases:

$$
\begin{array}{lll}
\Delta_{1,1}^m = \frac{m_4}{m_2} - 0.7 & \Delta_{2,1}^m = \frac{m_5}{m_2} - 0.5 & \Delta_{3,1}^m = \frac{m_6}{m_2} - 0.3 \\
\Delta_{1,2}^m = \frac{m_7 - m_2}{1 - m_2} - 0.7 & \Delta_{2,2}^m = \frac{m_8 - m_2}{1 - m_2} - 0.5 & \Delta_{3,2}^m = \frac{m_9 - m_2}{1 - m_2} - 0.3 \\
\Delta_{1,1}^h = h_4 - 0.7 & \Delta_{2,1}^h = h_5 - 0.5 & \Delta_{3,1}^h = h_6 - 0.3 \\
\Delta_{1,2}^h = \frac{h_7 - h_1}{h_3 - h_1} - 0.7 & \Delta_{2,2}^h = \frac{h_8 - h_1}{h_3 - h_1} - 0.5 & \Delta_{3,2}^h = \frac{h_9 - h_1}{h_3 - h_1} - 0.3.
\end{array}
$$

The four rows in this table correspond to the following triples of lotteries, respectively: lotteries with low money prizes; lotteries with high money prizes; lotteries with 'death' and 'complete health' as consequences; lotteries with the better and worse health states as consequences.

(d) Test on differences between probability distortions for money prizes and health states. The null hypotheses 'no difference', and, thus, independence of probability distortions on the nature of the prizes involved, is tested, for each of the three probabilities under consideration, against the alternatives that there is more distortion downwards in the case of health states. In other words, the alternatives reflect our intuition that individuals will be more risk averse if the risk concerns their health.

Formally, since we have two measurements for each probability distortion, we consider the mean values $\mu^m(0.7) = (1/2)(m_4/m_2) + (1/2)(m_7 - m_2)/(1 - m_2)$, $\mu^h(0.7) = (1/2)h_4 + (1/2)(h_7 - h_1)/(h_3 - h_1)$, and $\mu^m(0.5)$, $\mu^h(0.5)$, $\mu^m(0.3)$, and $\mu^h(0.3)$ defined analogously. Let $M_1 = \mu^m(0.7) - \mu^h(0.7)$, $M_2 = \mu^m(0.5) - \mu^h(0.5)$, $M_3 = \mu^m(0.3) - \mu^h(0.3)$. We test the null hypotheses $M_i = 0$ against the alternatives $M_i > 0$ $(i = 1, 2, 3)$.

The tests in (a) and (b) give some information on the consistency with which the subjects parttaking in the experiment distort probabilities in two separate cases, namely the case of monetary prizes and the case of health prizes. So actually, we test whether such distortions might depend on the height of the monetary prizes, or, alternatively, on the nature of the health states, thereby presuming that distortion functions may be different between monetary prizes and health 'prizes'. It is not hard to see that there will be some interdependencies beteen the tests in (a) and (b) on the one hand, and (c) on the other hand, in the sense that not all combinations of results will be possible. Just by way of an example, large differences in (a) exclude overall risk neutrality in (c). The tests in (c), on risk aversion, are of relevance mainly because they may provide evidence for or against expected utility theory. At their turn, the results in (c) are interdependent with the results of the tests in (d).

In (d), we test whether subjects distort probabilities differently, depending on the nature of the prizes. In this sense, in (d) we really test the validity of the rank-order approach as an overall theory. The results, presented in the next section, indicate that indeed — as we would expect intuitively — the distortion functions depend on the nature of the prizes involved, but not to the extent we expected before the experiment. Even, however, if these distortions would have turned out to be quite different, we would not consider this as evidence against the rank-order approach, since we do not believe that all uncertainty decisions of an individual can be described by just one simple function.

Practical set-up of the experiment

An interactive computer program, written in Pascal 5.0 on an IBM Personal System computer, is used to perform the experiment. Each subject runs the program in complete privacy; only an assistant is present (in a different room), to whom questions about the program can be directed. The subject first has to read a manual of three short pages, containing some general information, some information (examples) about the meaning of probabilities and lotteries, and the information needed to be able to answer the questions about health preferences. In particular, the three health conditions are described (see the Appendix to this report).

Next, the subject has to answer the questions formulated by the program. (The manual always remains available.) The program consists of four parts. In the first part, some personal questions have to be answered (age, sexe, occupation, and a few questions in relation to the health conditions under consideration). This part is intended mainly for future use of the program.

The second and third parts are identical: they contain the questions about money preferences. The second part is a dummy part, meant to provide the subject with some practical exercise. All questions are answered by producing - with the aid of the cursor keys - a coloured bar on the screen the length of which relates to the intended numerical values. A question is answered by confirming the length of this bar by pressing the return-key, but a number of possibilities to revise one's answer are built in. Also some consistency constraints are included: $0 \leq m_1 \leq m_2 \leq m_3 \leq 1$, $0 \leq m_6 \leq m_5 \leq m_4 \leq m_2 \leq m_9 \leq m_8 \leq m_7 \leq 1$. The program only accepts answers satisfying these constraints. The answers given with respect to the sure events (m_1, m_2, m_3) and to the previous lotteries in the same series remain always projected on the screen. Similar remarks hold about the fourth part of the program, which contains the questions about health conditions. In particular, also there consistency constraints are included:
$0 \leq h_1 \leq h_2 \leq h_3 \leq 1$, $0 \leq h_6 \leq h_5 \leq h_4 \leq 1$, $h_1 \leq h_9 \leq h_8 \leq h_7 \leq h_3$.
Before starting with the fourth parth, subjects are urged to read (anew) the descriptions of the three health states H_1, H_2, and H_3.

It turns out that, on the average, subjects need between 30 and 45 minutes to complete the whole procedure.

4 Results and Conclusions

Eighteen subjects participated in the experiment. These subjects were mainly undergraduate students in economics or health science. As noted before, their valuations were used in a Wilcoxon signed rank test.

For each of the sets of tests (a),(b),(c), and (d), we first list the results, and then draw a few conclusions, if any.

(a) Tests on consistency of probability distortions in case of monetary prizes

Results:
None of the null hypotheses could be rejected (level .10, two-tailed). Slight tendency towards positive values of Δ_i^m, $i = 1,2,3$.

Conclusions:
These results are in line with the rank-order approach. The mentioned tendency indicates a (non-significant) tendency towards more risk aversion (relatively speaking) in case of higher prizes.

(b) Tests on consistency of probability distortions in case of health states as prizes

Results:
For the probability 0.7, the null hypothesis, $\Delta_1^h = 0$, had to be rejected (level .01, two-tailed); the alternative hypothesis $\Delta_1^h < 0$ was accepted (level .005, one-tailed).
Also for the probability 0.5, the null hypothesis, $\Delta_2^h = 0$ had to be rejected (level just above .02, two-tailed); the alternative hypothesis $\Delta_2^h < 0$ was accepted (level .01, one-tailed)
For the probability 0.3, however, the null hypothesis, $\Delta_3^h = 0$, could not be rejected (level 0.1, two-tailed); there was only a very light tendency towards negative values of Δ_3^h.

Conclusions:
These results show that, in case of health states as prizes, the extent to which individuals distort probabilities depends very much on the 'height' of the prizes: they are significantly more risk averse (in a relative sense again) if prizes are 'death' and 'complete health', and if the probabilities of reaching the better state are relatively high (.7 and .5). However, the difference in probability distortion is less, and becomes non-significant, if there is a low probability (.3) of reaching the better state.

(c) Tests on risk aversion

Results:
The null hypotheses $\Delta_{i,1}^m = 0$ could not be rejected ($i = 1,2,3$, level .1, two-tailed).
The null hypotheses $\Delta_{i,2}^m = 0$ could be rejected for $i = 2,3$ with a level of .05, two-tailed; the corresponding alternatives could be accepted (levels slightly higher than .025, one-tailed). For $i = 1$, the alternative hypothesis could also be accepted, albeit with a higher level (.1, one-tailed).
The alternative hypotheses $\Delta_{i,1}^h < 0$ ($i = 1,2,3$) were all accepted (levels 0.005).
The null hypotheses $\Delta_{i,2}^h = 0$ could not be rejected for $i = 1,2$ (levels higher than .1). For $i = 3$, the alternative hypothesis could be accepted (level .025).

Conclusions:
As to the money lotteries, we found risk aversion only for the high prizes; for the low prize lotteries, we found risk neutrality, so actually consistency with expected utility theory. These results are only partly consistent with the results found under (a), since risk aversion for the high prize lotteries and risk neutrality for the low prize lotteries would seem to lead to rejection of the null hypotheses under (a). However, we did find a tendency towards relatively more risk aversion in case of higher prizes, in (a).
For the lotteries involving health states as prizes, we found risk aversion in all but two cases, namely those where there were relatively high probabilities (.7 and .5) of reaching the better health state, and death was not involved. We note that these results are consistent with the results found under (b).

(d) Comparison of probability distortions between monetary prizes and health states

Results:
The alternative hypothesis could only convincingly be accepted in one case, namely $M_3 > 0$ (level .025). The alternative hypothesis $M_2 > 0$ could be accepted with a level of .1, whereas the null hypothesis $M_1 = 0$ could not be rejected with a reasonable level, although there was a tendency towards $M_1 > 0$.

Conclusions:
Individuals tend to be relatively more risk averse if health states are concerned, and if there are relatively high probabilities (.7 and .5) of 'bad' prizes (low money prizes, bad health states). However, if there is a high probability (.7) of a 'good' prize and a correspondingly low probability of a 'bad' outcome, then this difference is no longer significant. This actually means that, in the lower range of its domain, the probability distortion function might be relatively independent of the nature of the 'prizes'.

In view of the preliminary character and limited set-up of the experiment, we should not attach universal value to the conclusions formulated above. We can, however, say a few things.

We found that the extent to which individuals tended to distort objective probabilities, depended on the 'height of the prizes' in case of health states, but not so much in case of money. Loosely speaking, individuals exhibited one distortion function for money lotteries, but two such functions for health state lotteries. Further, they turned out to be more risk averse in case of health. For low money lotteries, they even were risk neutral, so behaving in line with Von Neumann-Morgenstern (expected) utility theory. Finally, there was some evidence that the distortion function may be relatively independent of the nature of the prizes in the lower part of its domain. Note that the extent to which one could use these results for or against the rank-order approach, depends on how much one would want an individual probability distortion function to be sensitive to the nature of the 'prizes'.

One (not too surprising) conclusion we may draw from our experiment is that the standard gamble method for measuring utilities for health states is to be used with great care: most of our results for health state lotteries (in part (c)) indicate risk aversion according to the rank-order approach (and hence, violation of the Von Neumann-Morgenstern theory on which the standard gamble method is based), especially for lotteries leading to death or complete health. The results are better — that is, more in line with the Von Neumann-Morgenstern theory — if less risky events (i.c., involving health states 1 and 3) are taken to be the prizes: recall that there we only found statistical proof for risk aversion if there was only a low probability of reaching the better health state. This suggests that we should avoid using the SG method in case there is a possibility of a very bad outcome or a high probability of the worse of two outcomes.

It is important to note that, even if we would believe in the rank-order approach to utility, this does not mean that we would have a new method for measuring utility indices for health states. We still have to rely on methods like Visual Analogue or Time Trade-off to measure such utility indices (cf. (1)). Only in case we would like to have utilities for *risky* events - which is *not* the standard practice - would we have to resort to a method like Standard Gamble (cf. (2)). Otherwise, in case of riskless events, we may only rely on the SG method if we are convinced that individuals are risk neutral according to the rank-order theory or, equivalently, are expected utility maximizers according to the Von Neumann-Morgenstern theory, or, equivalently, φ is identity (in (1)).

The experiment reported in this paper was inspired partly by the fact that within the department of the second author a scenario study is being undertaken on the effects of Ischaemic Heart Disease Prevention Programs in the Netherlands. The descriptions of the health states used in the experiment (see the appendix to this paper) refer to diseases of the heart. Of course, our experiment is too preliminary and limited to serve as a database for determining utility indices and qaly's. Verweij et al. (1988) provide a thorough procedure for the evaluation of health states.

We intend to use the above results and preliminary conclusions in order to improve our experimental set-up and to organize a more rigorous and larger session in the near future. By this report, the least we hope to achieve is to draw attention to alternative theories that can serve as a basis for utility measurement, in particular in health science.

We summarize the main conclusions. First, postulating the existence of an individual probability distortion function φ as in (2), we found evidence that at least for low probabilities this function might be reasonably independent of prizes. More precisely, the average individual differences between valuations of money and health states were, especially in case of 'bad' lotteries, with low probabilities of the good outcomes, not as large as we had expected. In other words, the rank-order approach might be accurate in the lower range of probabilities. Second, risk aversion according to the rank-order approach was more apparent in this region, that is, for 'bad' lotteries. For 'good' lotteries (high probabilities of the better prizes) and low-money lotteries we found no conclusive evidence against the Von Neumann-Morgenstern theory.

Appendix

This appendix contains the descriptions of the three health states occurring in the experiment. These three states were composed with the aid of classifications devised by the New York Heart Association and the Canadian Cardiovascular Society, and by Goldman et al. (1981).

Health state 1

Every physical activity and every emotion cause a longlasting and severe pain and tightness of the chest. These symptoms may occur even in a state of complete rest. Also involuntary physical activity like digestion (especially immediately following meals) leads to severe pains and tightness of the chest.

Consequently, even in a state of as much rest as possible, severe pain and tightness occur very often each day.

Health state 2

In a state of complete rest you feel reasonably well. You can also just manage climbing one flight of stairs, or going for a not too long and quiet walk (walk around a small block). More than this, however, is impossible without causing severe pain and tightness of the chest. There is a serious limitation of ordinary physical activity.

Health state 3

Ordinary physical activity is slightly limited. Climbing stairways or hills at a more than normal pace, especially right after meals, leads to undue fatigue, shortness of breath, or even pain in the chest, and too many emotions may cause similar symptoms. If you bear this in mind, however, you may lead a normal life.

References

Allais, M. (1979): "The Foundations of a Positive Theory of Choice Involving Risk and a Criticism of the Postulates and Axioms of the American School", in: *Expected Utility and the Allais Paradox* (eds. M. Allais and O. Hagen), 27–145, Reidel, Dordrecht

Bombardier, C., A.D. Wolfson, A.J. Sinclair, and A. McGreer (1982): "Comparison of Three Preference Measurement Methodologies in the Evaluation of a Functional Status Index", in: *Choices in Health Care: Decision Making and Evaluation of Effectiveness* (eds. R. Deber and G. Thompson), University of Toronto

Chew, S.H., E. Karni, and Z. Safra (1987):"Risk Aversion in the Theory of Expected Utility with Rank Dependent Probabilities", *Journal of Economic Theory*, 42, 370-381

Gilboa, I. (1987): "Expected Utility with Purely Subjective Non-Additive Probabilities", *Journal of Mathematical Economics*, 16, 65-88

Goldman, L., H. Beverly, E. Francis Cook, and A. Loscalzo (1981): "Comparative Reproducibility and Validity of Systems for Assessing Cardiovascular Functional Class: Advantages of a New Specific Activity Scale", *Circulation*, 64, 1227-1234

Herstein, I.N., and J.W. Milnor (1953): "An Axiomatic Approach to Measurable Utility", *Econometrica*, 21, 291–297

Jaffray, J.Y. (1986): "Some Experimental Findings on Decision Making Under Risk and their Implications", mimeo, Université de Paris, Laboratoire d'Econometrie

Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision Under Risk", *Econometrica*, 47, 263–291

Lefoll, J. , J.M. Guiot, and A. Chateauneuf (1988): "Allais' Model vs. Expected Utility Model: Some Preliminary Empirical Results", paper presented at FUR IV conference, Budapest

Llewellyn-Thomas, H., H.J. Sutherland, R. Tibshirani, A. Ciampi, J.E. Till, and N.F. Boyd (1984): "Describing Health States: Methodological Issues in Obtaining Values for Health States", *Medical Care*, 22, 543–552

Loomes, G. (1988): "Disparities Between Health State Measures: An Explanation and Some Implications", mimeo, Dept. of Economics and Related Studies, University of York

Loomes G., and R. Sugden (1982): "Regret Theory: An Alternative Theory of Rational Choice Under Uncertainty", *Economic Journal*, 92, 805–824

Machina, M., (1987): "Choice Under Uncertainty: Problems Solved and Unsolved", *Economic Perspectives*, 1, 121–154

Peters, H., (1988): "A Criterion for Comparing Strength of Preference, with an Application", mimeo, University of Limburg, Maastricht, The Netherlands

Pliskin, J.S. , D.S. Shepard, and M.C. Weinstein (1980): "Utility Functions for Life Years and Health Status", *Operations Research*, 28, 206–224

Quiggin, J. (1982): "A Theory of Anticipated Utility", *Journal of Economic Behavior and Organization*, 3, 323–343

Schmeidler, D. (1984): "Subjective Probability and Expected Utility without Additivity", Caress working paper 84-21, University of Pennsylvania

Torrance, G.W. (1976): "Social Preferences for Health States: An Empirical Evaluation of Three Measurement Techniques", *Socio-Economic Planning Sciences*, 10, 128–136

Torrance, G.W. (1986): "Measurement of Health State Utilities for Economic Appraisal", *Journal of Health Economics*, 5, 1–30

Torrance, G.W., M.H. Boyle, and S.P. Horwood (1982): "Application of Multi-Attribute Utility Theory to Measure Social Preferences for Health States", *Operations Research*, 30, 1043-1069

Verweij, M.J., W.E. Saris, G. Bonsel, B. van Hont, J.D.F. Habbema, P.J. van der Maas, and F. Th. de Charro (1988): "A Measurement Procedure for the Evaluation of Health States by the General Public", *Kwantitatieve Methoden*, 29, 93–113

Von Neumann, J., and O. Morgenstern (1947): *Theory of Games and Economic Behaviour*, 2nd edition, Princeton University Press, Princeton

Wakker, P. (1988): "Transforming Probabilities Without Violating Stochastic Dominance", Internal Report 88 MA 01, Department of Mathematical Psychology, University of Nijmegen

Wakker, P. (1989): "Continuous Subjective Expected Utility with Nonadditive Probabilities", *Journal of Mathematical Economics*, 18, 1–27

Yaari, M. (1987): "The Dual Theory of Choice Under Risk", *Econometrica*, 55, 95–115

Ontvangen: 23-01-1990
Geaccepteerd: 05-02-1990