The Journal of Risk and Uncertainty, 23:2; 185–198, 2001

Probability Weighting in Choice under Risk: An Empirical Test

HAN BLEICHRODT* Bleichrodt@bmg.eur.nl
iMTA/iBMG, Erasmus University, Rotterdam, The Netherlands

Abstract

This paper reports a violation of rank-dependent utility with inverse S-shaped probability weighting for binary gambles. The paper starts with a violation of expected utility theory: one-stage gambles elicit systematically different utilities than theoretically equivalent two-stage gambles. This systematic disparity does not disappear, but becomes more pronounced after correction for inverse S-shaped probability weighting. The data are also inconsistent with configural weight theory and Machina's fanning out hypothesis. Possible explanations for the data are loss aversion and anchoring and insufficient adjustment.

Keywords: nonexpected utility, probability weighting, health

JEL Classification: D8, I10

Expected utility theory is characterized by three types of preference conditions (Jensen, 1967): an elementary rationality condition (weak ordering), a sophisticated rationality condition (von Neumann Morgenstern independence), and a technical condition (Jensen continuity). Most nonexpected utility theories have concentrated on the sophisticated rationality part of expected utility, weakening independence. Empirical research has displayed many violations of independence. An important cause of these violations is *probability weighting*, the nonlinear sensitivity of people towards probability. Probability weighting is an essential part of rank-dependent utility theory (Quiggin, 1982; Yaari, 1987) and its derivative cumulative prospect theory (Tversky and Kahneman, 1992), currently the most influential alternatives for expected utility. Several empirical studies have displayed the importance of probability weighting and have provided quantitative assessments of its effects (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992; Lattimore, Baker, and Witte, 1992; Camerer and Ho, 1994; Tversky and Fox, 1995; Wu and Gonzalez, 1996; Gonzalez and Wu, 1999; Bleichrodt, van Rijn, and Johannesson, 1999; Abdellaoui, 2000; Bleichrodt and Pinto, 2000). These studies are consistent with

* Address for correspondence: Han Bleichrodt, iMTA/iBMG, Erasmus University, PO Box 1738, 3000 DR Rotterdam, The Netherlands.

a probability weighting function that is inverse S-shaped: small probabilities are overweighted and intermediate and large probabilities are underweighted.

The aim of the present paper is to examine the contribution that rank-dependent utility with an inverse S-shaped probability weighting function (RDU_{IS}) can make to the explanation of violations of expected utility. The paper starts with a consistency test of expected utility: do utility elicitation procedures that are equivalent under expected utility lead to identical results? It is shown that individual preferences do not satisfy this consistency test and that a systematic bias occurs. This violation of expected utility has been observed before in Llewellyn-Thomas et al. (1982) and Rutten-van Mölken et al. (1995). The experiment reported in this paper differs in two respects from these previous studies. First, subjects were allowed to express imprecise preferences. Imprecise preferences were allowed to try and reduce the impact of response error on the results. Second, the experiment controlled for a framing bias. In spite of these experimental changes, the violation re-emerged. This indicates that the violation is robust.

The paper then examines whether RDU_{IS} can remove the systematic bias between the utilities. The conclusion is negative. Instead of vanishing, the systematic bias becomes more pronounced after correction for probability weighting. There exist other studies that report negative evidence on RDU_{IS} (Wakker, Erev, and Weber, 1994; Wu, 1994; Birnbaum and McIntosh, 1996; Birnbaum and Navarrete, 1998). These studies have in common that they involved gambles with at least three outcomes. In contrast, the present study involves only binary gambles. Luce (2000) argues on the basis of the available empirical evidence that the rank-dependent utility model is descriptively more valid for binary gambles than for gambles involving more than two outcomes. The new finding of the present paper is that rank dependent utility can be systematically violated for binary gambles. This finding is important for practical decision analysis because utility elicitation is generally based on binary gambles.

Birnbaum and McIntosh (1996) explained the violations of RDU_{IS} they observed by a configural weight model. The configural weight model cannot explain the data of this paper. For binary gambles the configural weight model yields predictions that are nearly identical to those of RDU_{IS}.

The violation of expected utility studied in this paper is different from the utility evaluation effect (Machina, 1983, 1987) observed by Karmarkar (1974, 1978), Allais (1979), and McCord and de Neufville (1983, 1984). As will be explained in the Discussion, the utility evaluation effect is consistent with RDU_{IS}. Machina (1983, 1987) showed that the utility evaluation effect can also be explained by Machina's (1982) Hypothesis II which implies that indifference curves fan out throughout the probability triangle. It will be shown that the violation of expected utility observed in this paper cannot be explained by fanning out of indifference curves and is therefore different than the utility evaluation effect.

The paper is organized as follows. Section 1 explains the theoretical differences between expected utility and rank-dependent utility. Section 2 describes the consistency test used to compare the two theories. Experimental procedures and results are described in Sections 3 and 4 respectively. Section 5 concludes.

1. Theory

The experimental questions, described in Section 3, only invoke binary gambles and therefore the formal analysis is restricted to such gambles. A typical *gamble* is denoted by $(x, p; y)$, yielding outcome x with probability p and outcome y with probability $1 - p$. The *outcomes* x, y are elements of a set of outcomes $\mathscr{X}$. In the experiment of this paper, the outcomes are health states. If $x = y$ then the gamble is *riskless*. Preferences over gambles are denoted by $\succcurlyeq$. Strict preferences are denoted by $\succ$ and indifferences by $\sim$. Preferences over outcomes coincide with preferences over riskless gambles. For notational convenience, it is assumed that gambles are *rank-ordered*, i.e. it is implicit in the notation $(x, p; y)$ that $x \succcurlyeq y$.

Under *expected utility*, preferences over gambles $(x, p; y)$ can be represented by:

$$p \cdot U(x) + (1 - p) \cdot U(y) \tag{1}$$

where U is a real-valued *utility function* defined on $\mathscr{X}$. U is unique up to positive affine transformations.

Under *rank-dependent utility*, preferences over gambles $(x, p; y)$ can be represented by

$$w(p) \cdot U(x) + (1 - w(p)) \cdot U(y) \tag{2}$$

where $w(\cdot)$ is a probability weighting function that has the following properties: $w(0) = 0$, $w(1) = 1$ and $w(p) \geq w(q)$ if and only if $p \geq q$. U is as before a real-valued function defined on $\mathscr{X}$ that is unique up to positive affine transformations. If the probability weighting function is linear, i.e., $w(p) = p$, for all p, then rank-dependent utility reduces to expected utility.

Empirical studies have displayed that the probability weighting function w is typically inverse S-shaped (Tversky and Kahneman, 1992; Camerer and Ho, 1994; Tversky and Fox, 1995; Wu and Gonzalez, 1996; Gonzalez and Wu, 1999; Bleichrodt, van Rijn, and Johannesson, 1999; Abdellaoui, 2000; Bleichrodt and Pinto, 2000). Tversky and Kahneman (1992) describe the probability weighting function by the following one-parametric specification:

$$w(p) = \frac{p^{\gamma}}{(p^{\gamma} + (1 - p)^{\gamma})^{\frac{1}{\gamma}}} \tag{3}$$

Table 1 gives an overview of the values of γ found in the literature. The empirical estimates are broadly similar. All estimates are consistent with an inverse S-shaped probability weighting function. The point where w changes from overweighting probabilities to underweighting probabilities, lies approximately at $p = 0.35$. These properties are confirmed in the non-parametric studies by Gonzalez and Wu (1999), Abdellaoui (2000), and Bleichrodt and Pinto (2000).

Table 1. Empirical estimates for the probability transformation parameter γ in Eq. (3)

Study	Monetary gains	Monetary losses	Health outcomes
Tversky and Kahneman (1992)	0.61	0.69	
Camerer and Ho (1994)	0.56		
Wu and Gonzalez (1996)	0.71		
Bleichrodt et al. (1999)			0.69
Abdellaoui (2000)	0.60	0.70	
Bleichrodt and Pinto (2000)			0.71

2. The consistency test

Consider the following medical decision problem. A client has to choose between two options: a health state x for certain and a risky treatment option (full health, p; immediate death). Suppose the client is indifferent between these two options for $p = 0.6$. By expected utility theory, this implies that

$$1 \cdot U(x) = 0.6 \cdot U(\text{full health}) + 0.4 \cdot U(\text{immediate death}) \tag{4}$$

Normalizing utility such that $U(\text{full health}) = 1$ and $U(\text{immediate death}) = 0$ then yields $U(x) = 0.6$.

Suppose further that the client is indifferent between health state y for certain and the treatment option (full health, 0.2; immediate death). That is,

$$1 \cdot U(y) = 0.2 \cdot U(\text{full health}) + 0.8 \cdot U(\text{immediate death}) \tag{5}$$

which implies by the above scaling that $U(y) = 0.2$.

The client is finally offered a choice between health state x for certain and the treatment option (full health, p'; health state y). Expected utility theory implies that he is indifferent between these two options if

$$U(x) = p' \cdot U(\text{full health}) + (1 - p') \cdot U(y) \tag{6}$$

$U(\text{full health}) = 1$, $U(x) = 0.6$ by Eq. (4), and $U(y) = 0.2$ by Eq. (5). Substituting these values into Eq. (6) gives $0.6 = p' \cdot 1 + (1 - p') \cdot 0.2$ and it follows that the client is consistent with expected utility if he states an indifference probability of 0.5.

Suppose now that the client is a rank-dependent utility maximizer and that his probability weighting function can be described by Eq. (3) with γ equal to 0.61, Tversky and Kahneman's value for gains. The indifference $x \sim$ (full health, 0.6; immediate death) implies:

$$U(x) = w(0.6) \cdot U(\text{full health}) + (1 - w(0.6)) \cdot U(\text{immediate death}) \tag{7}$$

Under the normalization U(full health) = 1 and U(immediate death) = 0, it follows that $U(x) = w(0.6) \approx 0.47$. A similar line of argument gives $U(y) = w(0.2) \approx 0.26$. Indifference holds in the third question if

$$U(x) = w(p') \cdot U(\text{full health}) + (1 - w(p')) \cdot U(y) \tag{8}$$

Substitution gives $0.47 = w(p') \cdot 1 + [1 - w(p')] \cdot 0.26$ or $w(p') = 0.284$ and $p' = w^{-1}(0.284) \approx 0.24$. Hence, a rank-dependent utility maximizer will state an indifference probability of 0.24 in the third choice question.

Throughout the paper, gambles in which the treatment outcomes are full health and immediate death, such as the first two gambles in the above example, are referred to as *one-stage gambles*. In a one-stage gamble, the utility of the certain outcome can be calculated directly given the normalization U(full health) = 1 and U(immediate death) = 0. Gambles in which death has been replaced as the worst outcome of treatment by a better health state y, such as the third gamble in the above example, are referred to as *two-stage gambles*. Two-stage gambles only allow calculation of the utility of the certain outcome after substitution of the utility of health state y. That is, the evaluation of a two-stage gamble requires the input of the response to a one-stage gamble involving health state y.

3. Experiment

Subjects and health states

Sixty-six health economics students at the Erasmus University participated in the experiment. The experiment was administered in 6 group sessions. The subjects faced 4 health states in addition to full health (health state F) and immediate death (health state D). The 4 health states are described in the Appendix. The health states were taken from the EuroQol classification system, a widely used method in health utility measurement (The EuroQol Group, 1990). Health states were described in terms of level of functioning on six attributes: mobility, self care, general daily activities, leisure activities, pain, and mental health. Full health was defined as no limitations on any of these attributes.

Tasks

Subjects were first asked to rank the health states in order of preference and subsequently to put the health states on a rating scale calibrated between 0 and 1. These tasks were included to familiarize subjects with the health states and to obtain information about their rank ordering of the health states. Subjects were then taken through three practice questions to familiarize them with the task of expressing preferences in terms of probabilities. The actual experiment consisted of 4 one-stage gambles and 3 two-stage

Table 2. The 7 questions

Number question	Certain outcome	Treatment outcomes	Number question	Certain outcome	Treatment outcomes
1.	A	F and D	5.	B	F and E
2.	B	F and D	6.	C	F and B
3.	C	F and D	7.	E	F and A
4.	E	F and D			

gambles. The gambles are described in Table 2. The order of presentation of the gambles was randomized to avoid order effects.

Subjects were encouraged to state first the probabilities for which they had a clear preference for either the certain outcome or the risky treatment $(x, p; y)$ and, finally, the probabilities for which they found it hard to choose between the two options. Subjects were not forced to state one indifference probability, because their preferences are typically somewhat imprecise (Dubourg, Jones-Lee, and Loomes, 1994). It was hoped that the impact of response error would be reduced by allowing subjects to express such imprecision and by analyzing the data for several probabilities throughout the range of probabilities for which subjects found it hard to choose.

Probability elicitation was by a line of values for the probability of successful treatment (p). Next to this line, a line was drawn that indicated the complementary probability of failure of treatment $(1 - p)$. This line was drawn to avoid a possible framing bias. The second line should remind subjects what a particular choice of p implied in terms of treatment failure. Only displaying the probability of successful treatment might induce subjects to overemphasize the outcome of successful treatment, ignoring the possibility of failure of treatment.

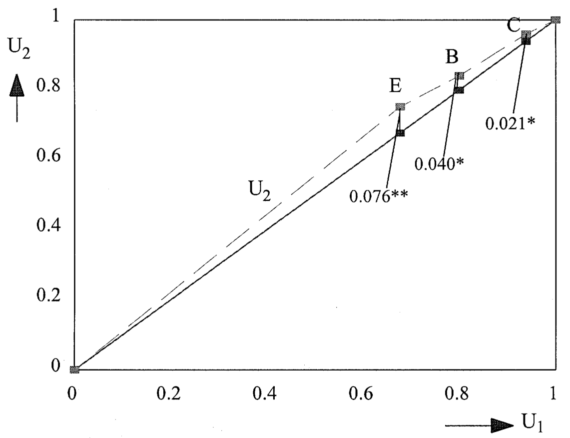
4. Results

Four subjects were excluded from the analyses because they failed to answer at least one question. Question 5 could only be analyzed for 50 subjects. The remaining 12 subjects ranked health state E above health state B. The other questions were analyzed for all 62 subjects. The results are similar if only the responses are used of the 50 subjects for whom question 5 could be analyzed.

The results were not affected by the probability that was selected from the range of probabilities for which subjects found it hard to choose between the options. The analyses reported in the paper are based on the midpoint of this range.

Expected utility

Figure 1 displays the utilities according to the one-stage gambles, U_1, the two-stage gambles, U_2, and their differences under expected utility. In violation of expected utility,



* Significantly different from zero at $\alpha = 0.05$

** Significantly different from zero at $\alpha = 0.01$

Figure 1. Comparison between one-stage (U_1) and two-stage (U_2) utilities under expected utility.

there is a systematic difference between U_1 and U_2: the two-stage utilities exceed the one-stage utilities. The difference is significant for all health states.

Rank-dependent utility

To compare one-stage and two-stage utilities under RDU_{IS}, Eq. (3) is adopted with $\gamma = 0.61$, Tversky and Kahneman's estimate for gains. As shown in Table 1, other empirical studies have found estimates for γ that differ only slightly from 0.61. The studies by Bleichrodt, van Rijn, and Johannesson (1999) and Bleichrodt and Pinto (2000) involved medical outcomes.

Figure 2 shows the comparison between one-stage and two-stage utilities under RDU_{IS}. Instead of vanishing or at least diminishing, the differences exacerbate after correction for probability weighting. The violations of rank-dependent utility persist if other values for γ that have been observed in the literature or other parametric specifications of the inverse S-shaped probability weighting function (Lattimore, Baker, and Witte, 1992; Prelec, 1998; Gonzalez and Wu 1999) are substituted. The divergence ceases to be systematic for $\gamma = 3.34$. This is an implausibly high value that contradicts all previous studies on probability weighting and that is inconsistent with an inverse S-shaped probability weighting function.

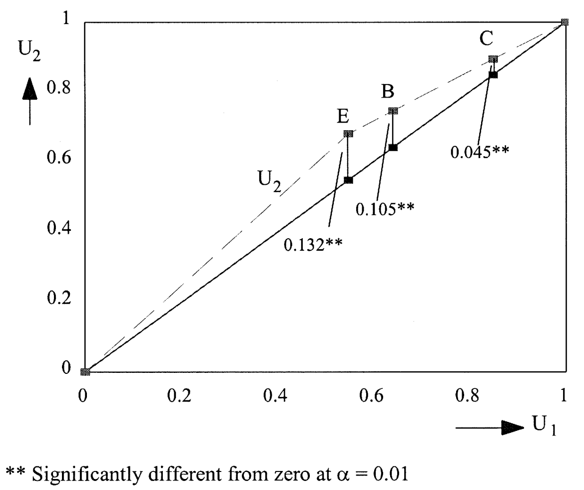


Figure 2. Comparison between one-stage (U_1) and two-stage (U_2) utilities under rank-dependent utility.

5. Discussion

Main finding

The main finding of this study is that RDU_{IS} is violated for binary gambles. Previous empirical studies found support both for rank-dependent utility for binary gambles and for the inverse S-shaped probability weighting function. However, this paper concludes that incorporating inverse S-shaped probability weighting in rank-dependent utility exacerbates rather than mitigates the systematic differences between one-stage and two-stage utilities that were observed under expected utility.

The utility evaluation effect

The paper displays a violation of expected utility in a probability equivalence task. Karmarkar (1974, 1978), Allais (1979), and McCord and de Neufville (1983, 1984) observed a violation of expected utility in a certainty equivalence task. They observed that the fractile method (Farquhar 1984) yields a more concave (expected) utility function when higher probabilities are used in the elicitation. Machina (1983, 1987) referred to this finding as the utility evaluation effect. Contrary to the violation of expected utility considered in this paper, the utility evaluation effect can be explained by RDU_{IS}. To illustrate, let x_q^p denote the outcome with expected utility q elicited with probability held fixed at value p throughout the certainty equivalence task. Table 3 shows the rank-dependent utilities of $x_{0.25}^p$, $x_{0.50}^p$, and $x_{0.75}^p$ for $p = \{0.10, 0.30, 0.50, 0.70, 0.90\}$ when the probability weighting function is equal to Eq. (3) with $\gamma = 0.61$. It is observed that $U(x_q^p) > U(x_q^{p'})$ iff $p < p'$. Hence, x_q^p decreases with p if utility is monotonic in

Table 3. Rank-dependent utility with inverse S-shaped probability weighting can explain the utility evaluation effect

		EU	
p	0.25	0.50	0.75
0.10	0.480	0.763	0.936
0.30	0.269	0.528	0.763
0.50	0.177	0.421	0.664
0.70	0.082	0.281	0.554
0.90	0.010	0.086	0.362

outcomes. x_q^p decreasing in p is equivalent to the (expected) utility function becoming more concave the higher the probability used in the fractile method.

Machina (1983, 1987) noted that the utility evaluation effect can be explained by his Hypothesis II (Machina, 1982) which implies that indifference curves fan out throughout the probability triangle. Figure 3 shows that fanning out of indifference curves cannot explain the observed systematic disparity between one-stage and two-stage gambles. The figure presents the example of Section 2 in the probability triangle. The vertical side of the triangle displays the probability of full health (F), the horizontal side the probability of immediate death (D), and the remaining probability is the probability of health state y. Northwest movements throughout the triangle represent increases in preference. The figure shows the prediction for the two-stage gamble when indifference curves fan out. Because the gamble (F, 0.6; D) lies on a lower indifference curve than the gamble (F, 0.5; y) fanning out predicts that the value of p reported in the second stage of the two-stage gamble will be lower than 0.5. The typical pattern that we observe is, however, that the reported value of p exceeds 0.5. The data of this paper could be explained by indifference curves that fan in. However, a theory of indifference curves that fan in is

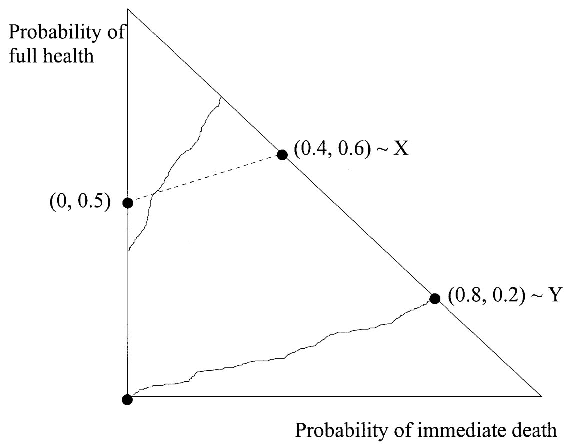


Figure 3. Fanning out of indifference curves cannot explain the data.

inconsistent with well-documented phenomena such as the common consequence effect and the common ratio effect.

Explanations

As noted in the Introduction, the data cannot be explained by configural weight theory either. One explanation for the data is suggested by cumulative prospect theory. Whereas the findings of this paper constitute a violation of rank-dependent utility, they do not necessarily violate cumulative prospect theory. Cumulative prospect theory generalizes expected utility by incorporating both probability weighting and *sign-dependence*, the dependence of preferences on the sign of outcomes. The experimental data of this paper were analyzed as if all outcomes were gains, i.e., as if subjects took death as their *reference point*. The results are not affected if all outcomes are interpreted as losses, i.e., if subjects took full health as their reference point. However, Hershey and Schoemaker (1985) conjectured that in answering probability equivalence questions of the type used in this paper, subjects take the certain outcome as their *reference point*. They then interpret the difference between the good gamble outcome and the certain outcome as a gain and the difference between the bad gamble outcome and the certain outcome as a loss. Appendix 2 gives a qualitative analysis of Hershey and Schoemaker's conjecture under cumulative prospect theory. A quantitative analysis is complicated. If Hershey and Schoemaker's conjecture is true then a subject's reference point varies between probability equivalence questions. A quantitative assessment of the predictions of cumulative prospect theory therefore requires the development of utility evaluation formulas that take into account variation in the reference point. The development of such formulas is beyond the scope of the present paper.

Appendix 2 shows that if Hershey and Schoemaker's conjecture is true then cumulative prospect theory with its usual assumptions of loss aversion and diminishing sensitivity (Tversky and Kahneman, 1992) can explain the systematic differences between one-stage and two-stage utilities observed under expected utility and rank-dependent utility. However, the Appendix also shows that under the reflection effect a model with only loss aversion would fit the data better than cumulative prospect theory. Since probability weighting is an important part of cumulative prospect theory, the data reported here are troubling for cumulative prospect theory as well.

An alternative explanation for the observed differences between one-stage and two-stage utilities is that the elementary rationality part of expected utility is violated. Previous research has argued that in responding to probability equivalence questions subjects tend to take one, the probability of the certain outcome, as an anchor and adjust their response downwards from one. This adjustment is generally insufficient leading to reported indifference probabilities that are too high and, hence, to utilities that are too concave (Tversky, Sattath, and Slovic, 1988; Johnson and Schkade, 1989). In the one-stage gambles this upward bias is present only once, namely in the reported indifference probability. However, in the two-stage gambles it is present twice. The double bias occurs because the utility of the certain outcome cannot be inferred directly but requires the input from another one-stage gamble, which in turn is biased upwards. Hence, in a two-stage gamble not only the reported indifference probability is biased upwards, but also the utility of the health state which was substituted for immediate death as the worst outcome of treatment. This additional upward bias can explain why two-stage utilities exceed one-stage utilities.

Concluding remarks

Several papers have argued that utilities should be adjusted for probability weighting (Fellner, 1961; Wakker and Stiggelbout, 1995) and the empirical literature suggests that this correction should be based on an inverse S-shaped probability weighting function. The findings of this paper challenge the generality of these arguments. They suggest that there are other factors besides probability weighting that affect people's responses to utility elicitation questions, that these factors may lead to biases opposite to the predictions of probability weighting, and that these biases may dominate the effect of probability weighting. Successful application of rank-dependent utility depends on the decision context, even for binary gambles.

Appendix 1: Description of the health states used in the experiment

A

1. Unable to walk without a stick, crutch or walking frame
2. No problems with self-care
3. Unable to perform main activity (work, study, housework)
4. Unable to pursue some family and leisure activities
5. Extreme pain or discomfort
6. Anxious or depressed

B

1. No problems in walking about
2. No problems with self-care
3. Unable to perform main activity (work, study, housework)
4. Able to pursue all family and leisure activities
5. Extreme pain or discomfort
6. Not anxious or depressed

C

1. No problems in walking about
2. No problems with self-care
3. Able to perform main activity (work, study, housework)
4. Able to pursue all family and leisure activities
5. Moderate pain or discomfort
6. Not anxious or depressed

E

1. No problems in walking about
2. No problems with self care
3. Unable to perform main activities (work, study, housework)
4. Unable to pursue some family and leisure activities
5. Moderate pain or discomfort
6. Anxious or depressed

Appendix 2: Explanation of the way in which loss aversion can account for the data

Let $F - x$ denote the gain of being in full health rather than in health state x, let $D - x$ denote the loss of being dead rather than in health state x, and let $y - x$ denote the loss of being in health state y rather than in health state x. Under Hershey and Schoemaker's (1985) conjecture and cumulative prospect theory, the one-stage gamble indifference $x \sim$ (full health, p; immediate death) yields; $U(x) = w^{+}(p) \cdot U(F - x) + w^{-}(1 - p) \cdot U(D - x)$. The two-stage gamble indifference $x \sim$ (full health, p'; y) yields $U(x) = w^{+}(p') \cdot U(F - x) + w^{-}(1 - p') \cdot U(y - x)$. Under loss aversion, the loss part has a relatively strong impact on people's preferences. Under diminishing sensitivity, $U(D - x)$ and $U(y - x)$ will not be too far apart. The stronger loss aversion and diminishing sensitivity, the closer p' will be to p. The fact that p' is relatively close to p causes the systematic bias between one-stage and two-stage utilities observed under expected utility and rank-dependent utility.

Even though sign dependence enfeebles some of the negative impact of probability weighting, the conclusion remains true that probability weighting as such exacerbates the disparities between one-stage and two-stage utilities. Most studies have found that $w^{-}(1 - p)$ does not differ much from $1 - w^{+}(p)$. This finding is referred to as the *reflection effect*. Under the reflection effect, the modelling of probability weighting in cumulative prospect theory is similar to that under rank-dependent utility theory and, hence, probability weighting aggravates the differences between one-stage and two-stage utilities. Therefore a model with only loss aversion will provide a better description of the data of this paper than cumulative prospect theory, which incorporates both loss aversion and probability weighting.

Acknowledgments

I am grateful to Peter Wakker and an anonymous referee for many helpful comments. This research was made possible by a fellowship from the Royal Netherlands Academy of Arts and Sciences.

References

Abdellaoui, M. (2000). "Parameter-Free Elicitation of Utilities and Probability Weighting Functions," *Management Science* 46, 1497–1512.

Allais, M. (1979). "The So-Called Allais Paradox and Rational Decisions under Uncertainty." In M. Allais and O. Hagen (eds.), *Expected Utility Hypotheses and the Allais Paradox*. Dordrecht: D. Reidel.

Birnbaum, M. H. and W. R. McIntosh. (1996). "Violations of Branch Independence in Choices between Gambles," *Organizational Behavior and Human Decision Processes* 67, 91–110.

Birnbaum, M. H. and J. B. Navarrete. (1998). "Testing Descriptive Utility Theories: Violations of Stochastic Dominance and Cumulative Independence," *Journal of Risk and Uncertainty* 17, 49–78.

Bleichrodt, H. and J.-L. Pinto. (2000). "A Parameter-Free Elicitation of the Probability Weighting Function in Medical Decision Analysis," *Management Science* 46, 1485–1496.

Bleichrodt, H., J. van Rijn, and M. Johannesson. (1999). "Probability Weighting and Utility Curvature in QALY Based Decision Making," *Journal of Mathematical Psychology* 43, 238–260.

Camerer, C. F. and T. -H. Ho. (1994). "Nonlinear Weighting of Probabilities and Violations of the Betweenness Axiom," *Journal of Risk and Uncertainty* 8, 167–196.

Dubourg, W. R., M. W. Jones-Lee, and G. Loomes. (1994). "Imprecise Preferences and the WTP-WTA Disparity," *Journal of Risk and Uncertainty* 9, 115–133.

The EuroQol Group. (1990). "EuroQol: A New Facility for the Measurement of Health Related Quality of Life," *Health Policy* 16, 199–208.

Farquhar, P. (1984). "Utility Assessment Methods," *Management Science* 30, 1283–1300.

Fellner, W. (1961). "Distortion of Subjective Probabilities as a Reaction to Uncertainty," *Quarterly Journal of Economics* 75, 670–689.

Gonzalez, R. and G. Wu. (1999). "On the Form of the Probability Weighting Function," *Cognitive Psychology* 38, 129–166.

Hershey, J. C. and P. J. H. Schoemaker. (1985). "Probability versus Certainty Equivalence Methods in Utility Measurement: Are They Equivalent?" *Management Science* 31, 1213–1231.

Jensen, N. E. (1967). "An Introduction to Bernoullian Utility Theory: I. Utility Functions," *Scandinavian Journal of Economics* 69, 163–183.

Johnson, E. J. and D. A. Schkade. (1989). "Bias in Utility Assessments: Further Evidence and Explanations," *Management Science* 35, 406–424.

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

Karmarkar, U. A. (1974). "The Effect of Probabilities on the Subjective Evaluation of Lotteries," MIT Working Paper No. 698–74, MIT, Cambridge, MA.

Karmarkar, U. A. (1978). "Subjectively Weighted Utility: A Descriptive Extension of the Expected Utility Model," *Organizational Behavior and Human Performance* 21, 61–72.

Lattimore, P. M., J. R. Baker, and A. D. Witte. (1992). "The Influence of Probability on Risky Choice," *Journal of Economic Behavior and Organization* 17, 377–400.

Llewellyn-Thomas, H., H. J. Sutherland, R. Tibshirani, A. Ciampi, J. E. Till, and N. F. Boyd. (1982). "The Measurement of Patients' Values in Medicine," *Medical Decision Making* 2, 449–462.

Luce, R. D. (2000). *Utility of Gains and Losses: Measurement-Theoretical and Experimental Approaches*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Machina, M. (1982). " 'Expected Utility' Analysis without the Independence Axiom," *Econometrica* 50, 277–323.

Machina, M. (1983). "Generalized Expected Utility Analysis and the Nature of Observed Violations of the Independence Axiom." In B. P. Stigum and F. Wenstop (eds.), *Foundations of Utility and Risk Theory with Applications*. Dordrecht: D. Reidel.

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

McCord, M. R. and R. de Neufville. (1983). "Empirical Demonstration that Expected Utility Decision Analysis Is Not Operational." In B. P. Stigum and F. Wenstop (eds.), *Foundations of Utility and Risk Theory with Applications*. Dordrecht: D. Reidel.

McCord, M. R. and R. de Neufville. (1984). "Utility Dependence on Probability: An Empirical Demonstration," *Journal of Large Scale Systems* 6, 91–103.

Prelec, D. (1998). "The Probability Weighting Function," *Econometrica* 66, 497–528.

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

Rutten-van Mölken, M. P., C. H. Bakker, E. K. A. van Doorslaer, and S. van der Linden. (1995). "Methodological Issues of Patient Utility Measurement. Experience from Two Clinical Trials," *Medical Care* 33, 922–937.

Tversky, A. and C. Fox. (1995). "Weighting Risk and Uncertainty," *Psychological Review* 102, 269–283.

Tversky, A. and D. Kahneman. (1992). "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* 5, 297–323.

Tversky, A. S. Sattath, and P. Slovic. (1988). “Contingent Weighting in Judgment and Choice,” *Psychological Review* 95, 371–384.

Wakker, P. P., I. Erev, and E. U. Weber. (1994). ”Comonotonic Independence: The Critical Test between Classical and Rank-Dependent Utility,” *Journal of Risk and Uncertainty* 9, 195–230.

Wakker, P. P. and A. M. Stiggelbout. (1995). “Explaining Distortions in Utility Elicitation Through the Rank-Dependent Model for Risky Choices,” *Medical Decision Making* 15, 180–186.

Wu, G. (1994). “An Empirical Test of Ordinal Independence,” *Journal of Risk and Uncertainty* 9, 39–60.

Wu, G. and R. Gonzalez. (1996). “Curvature of the Probability Weighting Function,” *Management Science* 42, 1676–1690.

Yaari, M. E. (1987). “The Dual Theory of Choice under Risk,” *Econometrica* 55, 95–115.