Decision Utility Theory: Back to von Neumann, Morgenstern, and Markowitz

Krzysztof Kontek

Artal Investments

1. December 2010

Decision Utility Theory: Back to von Neumann, Morgenstern, and Markowitz

Krzysztof Kontek[1,2]

Abstract

Prospect Theory (1979) and its Cumulative version (1992) argue for probability weighting to explain lottery choices. Decision Utility Theory presents an alternative solution, which makes no use of this concept. The new theory distinguishes decision and perception utility, postulates a double S-shaped decision utility curve similar to one hypothesized by Markowitz (1952), and applies the expected *decision* utility value similarly to the theory by von Neumann and Morgenstern (1944). Decision Utility Theory proposes straightforward risk measures, presents a simple explanation of risk attitudes by using the aspiration level concept, and predicts that people might not consider probabilities and outcomes jointly, on the contrary to the expected utility paradigm.

JEL classification: C91, D03, D81

Keywords: Expected Utility Theory, Markowitz Hypothesis, Prospect Theory, Decision Utility, Allais Paradox, Common Ratio Effect, Risk Attitude Measures, Aspiration Level.

1 Introduction

The concept of utility was first introduced by Bernoulli and Cramer in the early 18th century to solve the famous St. Petersburg Paradox. Both men assumed that the utility curve had to be concave in shape if it was to describe the decreasing marginal value of money. They also proposed the expected utility method of lottery valuation. Shortly afterwards, Bentham, Mill and Sidgwick proposed that people ought to desire those things that will maximize their utility (Read, 2004). Those early notions of *cardinal* utility, i.e. being able to describe the level of pleasure or satisfaction, were abandoned with time. Pareto was aware of the difficulty of deriving a utility

---

[1] Artal Investments, ul. Chroscickiego 93/105, 02-414 Warsaw, Poland, e-mail: kkontek2000@yahoo.com

[2] The author would like to thank several people for the discussions, comments or help, which have led to this paper: Harry Markowitz, Tomasz Berent, Konrad Rotuski, Tadeusz Tyszka and the participants of seminars, Joanna Sokołowska, Steve Canty, Jonathan Leland, and Michael Birnbaum.

function for goods (van Praag, 1999). Choosing between alternatives can, however, be described by an *ordinal* utility function. This will show that one utility is better than another, but not by how much. Robbins (1932) was the first to proclaim that utility was immeasurable, but this was soon challenged when von Neumann and Morgenstern published their *Theory of Games and Economic Behavior* (1944). The authors put forward an axiomatic approach; if an individual satisfies the rationality axioms then her preferences can be represented by a utility function u and the lotteries under consideration can be ranked according to their expected utility value:

$$v\left(\{x;p\}\right)=\sum_{i=1}^{n}u\left(x_i\right)p_i=E\left[u\left(x\right)\right] \tag{1.1}$$

where x and p are lottery outcomes and their respective probabilities of winning. Von Neumann and Morgenstern further suggested a method of measuring utility that used lotteries. Several problems have been encountered since then. First, determining utility using lotteries produces inconsistent results, especially when different probability values are applied. Second, the curve concavity typically assumed to describe risk aversion does not explain why people buy lottery tickets and insurances. This problem was the subject of intensive research by Friedman, Savage (1948) and Markowitz (1952), who proposed utility curves with convex segments in order to explain these phenomena. The most important problem, however, is that an increasing number of psychological experiments have pointed to the irrationality of human decision making, where "irrationality" is here understood to mean any deviation from the axioms underlying Expected Utility Theory. Most of these objections have concerned the independence axiom with the Allais Paradox being the main example of its violation. Utility has been rejected as a sufficient explanation for these paradoxes, and the concept of probability weighting has been introduced into the lottery valuation formula:

$$v\left(\{x;p\}\right)=\sum_{i=1}^{n}u\left(x_i\right)w\left(p_i\right) \tag{1.2}$$

This has led to the development of several theories commonly referred as Non-Expected Utility Theories. These include Prospect Theory (Kahneman, Tversky, 1979) and its Cumulative version (Tversky, Kahneman, 1992). The latter is much more complex than the former, as it incorporates cumulative probability representation to avoid the stochastic dominance violations introduced by probability weighting.

This paper introduces Decision Utility Theory, which offers an alternative approach to solve the problems described above. The theory distinguishes between perception and decision utility. The former is close in meaning to the utilities assumed by Bernoulli, Cramer, von Neumann and Morgenstern, and Kahneman and Tversky in that it states how a given amount of money is perceived by a subject. The latter is the utility which describes decisions made under conditions of risk[3]. The new theory postulates a double S-shaped decision utility curve similar to the one hypothesized by Markowitz (1952), and applies the expected *decision* utility valuation similarly to the theory derived by von Neumann and Morgenstern (1944).

2 Original von Neumann-Morgenstern Method

Von Neumann and Morgenstern proposed the following method of utility measurement in their classic *The Theory of Games and Economic Behavior* (1944[4]): "*Consider three events, C, A, B, for which the order of the individual's preferences is the one stated. Let p be a real number between 0 and 1, such that A is exactly equally desirable with combined event consisting of a chance of probability 1 - p for B and the remaining chance of probability p for C. Then we suggest the use of p as a numerical estimate for the ratio of the preference of A over B to that of C over B*" (p. 18, the notion of probability has been changed from 'α' in the original text to 'p'). Their suggestion may be presented as follows:

$$p = \frac{u(A) - u(B)}{u(C) - u(B)} \tag{2.1}$$

where u denotes the utility of an event. The reasoning behind (2.1) proceeds as follows. If utility were absent, event A would be the expected value of events B and C; and when expressed relatively within the [B, C] range, it would assume the value of probability p. Both values differ, however, in practice, which indicates the existence of a nonlinear utility. Determining the probability values p for a sequence of events A should thus enable utility to be measured.

The utility of A assumes values in the $\left[u(B),\, u(C)\right]$ range, as probability p assumes values between 0 and 1. Von Neumann and Morgenstern argued that "*utility is a number up to a*

[3] The term "decision utility" appears in some recent works by Kahneman (Kahneman, 1999, Kahneman, Wakker, Sarin, 1997, Kahneman, Thaler, 2006). Decision utility is defined as follows: *"The weight that is assigned to the desirability of an outcome in the context of a decision is called its decision utility. Decision utilities are inferred from choices and are used to explain choices"*. Kahneman claims that: *"The Prospect Theory value function represents the decision utility of the gains and losses associated with possible outcomes"*. This, however, is not correct. Although the value function can be inferred from choices, it cannot explain them unaided.

[4] The source of the citations is the third edition from 1953.

linear transformation" (p. 25). In fact (2.1) remains the same when utility u is substituted with $u_1 = a\,u + b$, for any b and $a \neq 0$. Therefore $u(B) = 0$ and $u(C) = 1$ can arbitrarily be assumed, in which case (2.1) simplifies to:

$$p = u(A) \tag{2.2}$$

It appears that the utility of event A is expressed by the probability of its occurrence. Greater the probability, greater is the event utility. Contrary to the common approach, the probability appears to be the utility, rather than a variable for the utility determination. This is quite understandable as it is the probability which decides choices between events in this case.

Let us now assume that the events B, A, and C are simply monetary outcomes x_1, x_2, and x_3, such that $x_1 < x_2 < x_3$, and that the function u' has been obtained by varying the value of x_2. This function is presented in Figure 2.1 (left) and has the concave shape frequently assumed by economists to describe risk aversion. Let us now consider another pair of outcomes x_4 and x_5 such that

$$x_1 < x_4 < x_5 < x_3 \tag{2.3}$$

Similarly to the previous case, utility can be determined for values x_2 such that

$$x_4 < x_2 < x_5 \tag{2.4}$$

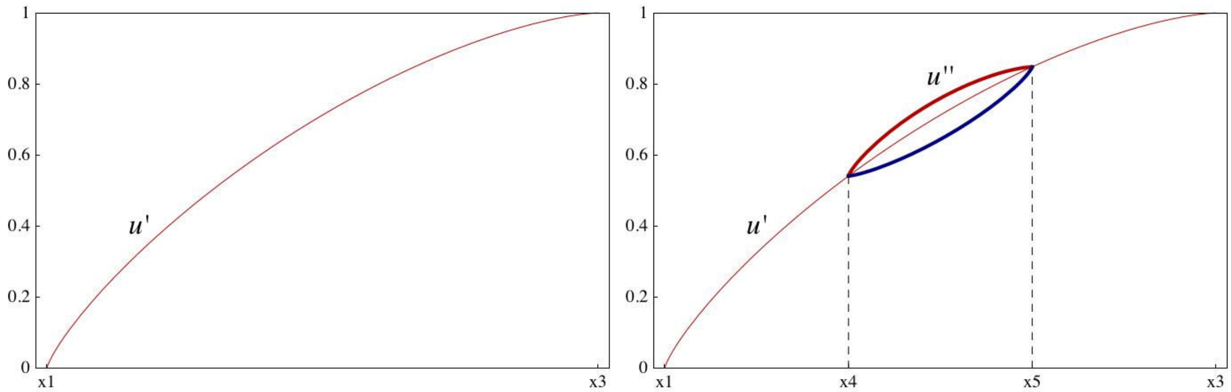

Figure 2.1 Left - Hypothetical utility *u'* in the range [x_1, x_3]. Right – Hypothetical utilities *u''* in the range [x_4, x_5] linearly transformed to have their endpoints located on *u'*.

The resulting utility u'' can be linearly transformed to have its endpoints located on utility u'. This means:

$$u''(x_4) = u'(x_4) \tag{2.5}$$

and

$$u''(x_5) = u'(x_5) \tag{2.6}$$

Hypothetical utilities *u''* are presented together with utility *u'* in Figure 2.1 (right). The question may be posed as to whether the utility curves for *u'* and *u''* have the same shape in the range [x_4, x_5]. This question may sound odd. There is no reason why their shapes should differ as a unique utility function defined over the outcome domain is implicitly assumed. Nor is there any reason why the method of determining utility, viz. over a wider or narrower range, should have any impact on the result. Despite these objections, it is assumed that utilities *u'* and *u''* may have different shapes in the range [x_4, x_5]. Instead of entering into a discourse on the soundness of this assumption, the experimental data will be presented in the next Point.

3 Experimentally Determined Utilities

A huge number of experiments on the shape of the utility curve have been conducted since von Neumann and Morgenstern's work. One of the best known are those of Tversky and Kahneman (1992), on the basis of which their Cumulative Prospect Theory was derived.

These experiments consisted of 28 lotteries for gains and 28 lotteries for losses. There were 7 pairs of lottery outcomes: {0, 50}, {0, 100}, {0, 200}, {0, 400}, {50, 100}, {50, 150}, and {100, 200}. The amounts were given in $ and defined the lottery outcome ranges. Tversky and Kahneman determined certainty equivalents for several probabilities within each range. Determining equivalents is the reverse of the method presented in Point 2, viz. to determining probabilities for given outcome values. It may, however, be assumed that people give consistent responses and that this reversed approach does not impact the results. The experimental data enable the utility functions to be plotted for each lottery range (see Figure 3.1). The plots for loss lotteries are not presented here, but the results are similar to those for gains.

There are at least two surprising findings in the plots obtained. First, the utilities are not concave as usually assumed by economists and by Tversky and Kahneman themselves. Instead, they are S-shaped and, interestingly, seem to be similar in all the ranges considered.

Second, the shape of the utility curve for a given outcome depends on the range being considered. For instance, utility is concave in the range [0, 50] for an outcome of $40, but roughly linear in the range [0, 100] for the same outcome value. More striking results can be observed when the outcome is $75. Utility is concave in the range [0, 100], roughly linear in the ranges [0, 200] and [50, 100] and convex in the ranges [0, 400] and [50, 150]. Similar differences can be observed for an outcome of $125. Utilities are concave in the ranges [0, 200] and [50, 150], but convex in the ranges [0, 400] and [100, 200]. It looks like a single utility working in all lottery ranges is not present.

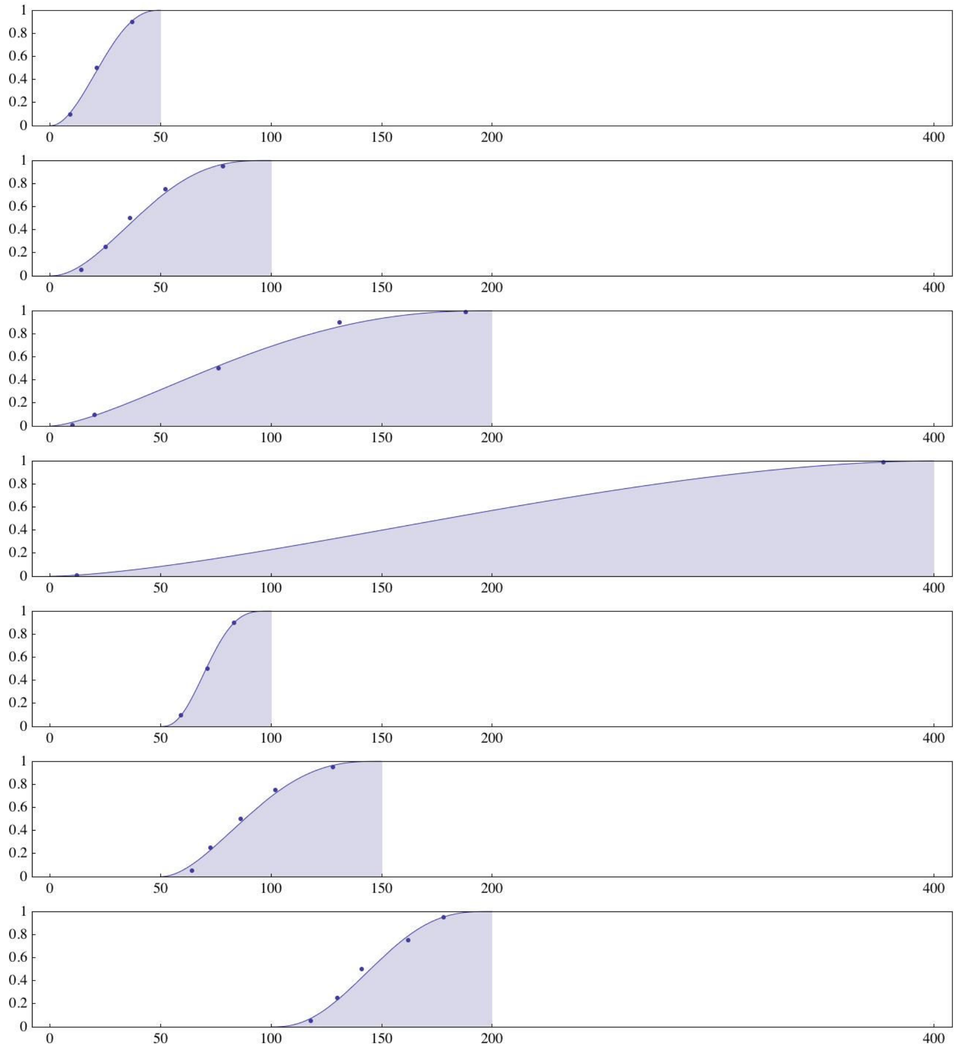

Figure 3.1 Utilities obtained using Tversky and Kahneman's data (1992). Each plot corresponds with the respective lottery ranges: [0, 50], [0, 100], [0, 200], [0, 400], [50, 100], [50, 150], and [100, 200].

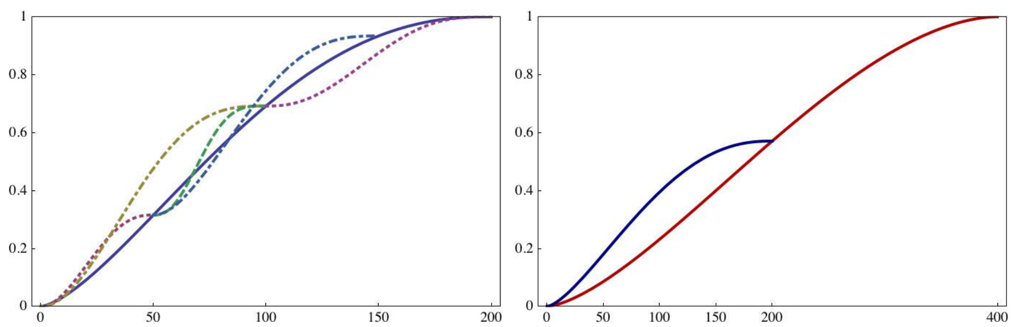

Figure 3.2 Left – Utilities in the range [0, 50], [0, 100], [50, 100], [50, 150], and [100, 200] linearly transformed to have their endpoints on the utility in the range [0, 200]. Right – The utility in the range [0, 200] linearly transformed to have its endpoints on the utility in the range [0, 400]

The shapes are analyzed in greater detail in Figure 3.2 (left). The utility in the range [0, 200] serves here as the pattern. Utilities in the ranges [0, 50], [0, 100], [50, 100], [50, 150], and [100, 200] are linearly transformed to have their endpoints located on the pattern utility. Very clearly, they do not match the shape of the pattern. What's more, the utility in the range [0, 200] does not match the utility in the range [0, 400] (Figure 3.2 - right) at all.

The presented results show that the utilities in each range do not match each other and that it is impossible to determine a single utility that would describe the experiments in all the sub-ranges. The problem with the lottery description therefore begins with the fundamental assumption that a single utility may be applied to lotteries having different outcome ranges. It follows that violations and paradoxes may appear when a single utility is used by a descriptive model. It follows also that the probability weights introduced by Prospect Theory serve merely to correct inexactnesses introduced by this utility. The solution of the problem can, however, be found using a different approach than by assuming the probability weighting concept.

4 Decision Utility

As stated in Point 3, the utility curves in the different lottery ranges seem to have a similar shape. This may lead to the hypothesis that the utility may be described within the range. In order to test this hypothesis, all the outcomes and certainty equivalents are linearly transformed to the [0, 1] interval using:

$$r = \frac{ce - P_{min}}{P_{max} - P_{min}}, \tag{4.1}$$

where r denotes the relative (normalized) outcome, ce denotes the certainty equivalent, $P_{max} = \mathrm{Max}(x)$ is the maximum lottery outcome, and $P_{min} = Min(x)$ is the minimum lottery outcome. All relative outcomes r together with their respective probabilities p are then presented on a single graph (Figure 4.1). All the points plotted very clearly create an S-shape, which can easily be estimated[5]. It turns out that a decision utility function D defined as:

$$p = D(r) \tag{4.2}$$

can be determined for outcomes expressed in relative terms.

[5] The Cumulative Beta Distribution has been used for this purpose, as it is the best known and most widely used distribution defined over the interval [0, 1].

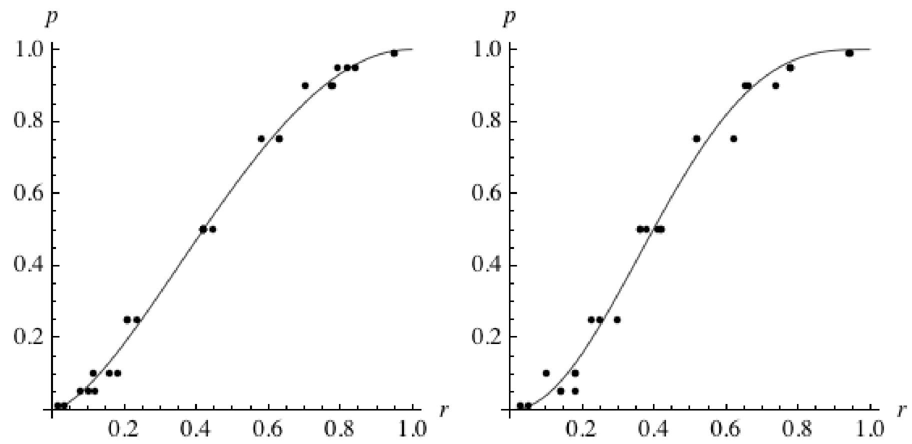

Figure 4.1 The decision utility *p* = *D*(*r*) determined using Tversky and Kahneman's data (1992). Left – the curve for losses. Right – the curve for gains. Points represent probabilities *p* for respective relative (normalized) outcomes.

Decision utility is very much different from the kind of utility assumed by von Neumann and Morgenstern, although it has been derived using their original method. It focuses on the description of decisions made under conditions of risk, rather than on how people perceive different welfare or income levels. Please note that decisions are always framed according to the model derived, as decision utility is defined in the [0, 1] interval. In fact transformation (4.1) is nothing more than problem framing[6]. However, as a result, the valuation of lotteries in absolute terms requires some further steps. The decision utility has to be transformed to the relative certainty equivalent:

$$ce_r = D^{-1}(p) \tag{4.3}$$

which then has to be rescaled to its absolute value:

$$ce = P_{min} + ce_r\left(P_{max} - P_{min}\right) \tag{4.4}$$

For $P_{min} = 0$, (4.4) simplifies to:

$$ce = P_{max}\, D^{-1}(p) \tag{4.5}$$

The procedure is very similar in the case of multi-outcome lotteries except that the (equivalent) probability has to be calculated first (refer Point 9). Lotteries can thus be compared by their certainty equivalents rather than by hypothetical "utils" as postulated by Expected Utility and Prospect Theory. This is a much more intuitive method, which additionally allows for cardi-

[6] This notion of framing is slightly different than the one normally discussed in the Prospect Theory context. In the latter case, framing is usually understood as considering the problem as a prospective gain or loss. In the present case, the problem is bounded from both sides, which is closer to the dictionary definition of frames.

nal measurements instead of simple option ordering.

5 Common Ratio Effect

As an example, the Common Ratio Effect analyzed by Kahneman and Tversky (1979) can now be explained using the decision utility model.

Problem 1: Choose between

A: 4,000 with a probability of 0.80 or B: 3,000 with a probability of 1.00

0 with a probability of 0.20

Problem 2: Choose between

C: 4,000 with a probability of 0.20 or D: 3,000 with a probability of 0.25

0 with a probability of 0.80 0 with a probability of 0.75

Experimental results consistently reveal that most people choose option B in Problem 1 and option C in Problem 2. Expected Utility Theory, per contra, predicts that people would choose either (A and C) or (B and D), as the probabilities of winning the main prize in the second pair of choices differ by a common ratio factor of 4 compared with the first pair. These options are presented on the utility axis (see Figure 5.1, left).

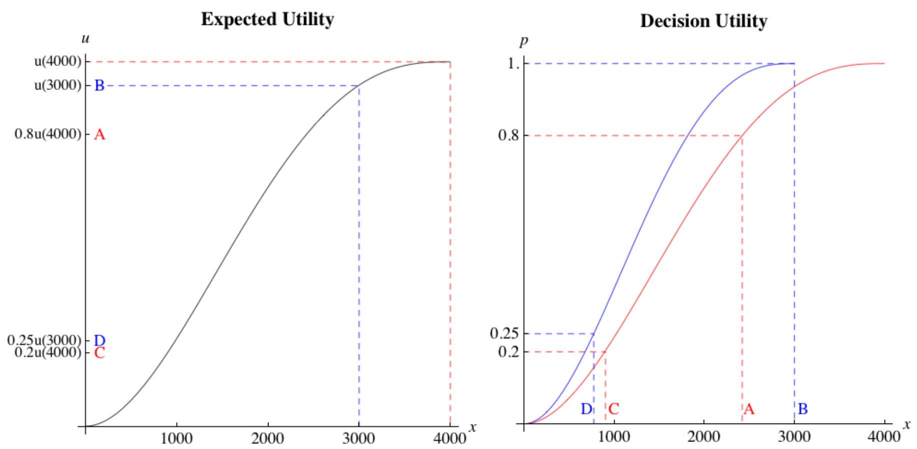

Figure 5.1 The Common Ration Effect presented as paradox using Expected Utility (left) and explained using the decision utility function (right).

The solution of this problem nicely demonstrates how a single decision utility applied to different lottery ranges predicts inconsistent choices (see Figure 5.1, right). The shape of both utility curves is the same as they are rescaled decision utility functions (cf (4.5)). The range of the red curve is [0, 4000] as this corresponds with options A and C, which have a maximum outcome of 4,000. The range of the blue curve is [0, 3000] as this corresponds with options B and D, which have a maximum outcome of 3,000. The certainty equivalents of each option can be found on the x-axis as presented in the figure. For greater probabilities, option B prevails over option A, but for lower probabilities, C is a better option than D. This is precisely what is observed in experiments. Importantly, the solution presented requires an S-shaped utility curve. A concave one, as presented in Figure 2.1, would generally not work in this case.

6 Decision Utility and the Markowitz Hypothesis

The decision utility for both gains and losses is jointly plotted using shapes presented in Figure 4.1. This utility is presented in Figure 6.1. Quite surprisingly, its shape resembles the one hypothesized by Markowitz (1952). It has three inflection points right where Markowitz predicted they would be. The function is monotonically increasing and is bounded from the top and the bottom. Concavities and convexities occur in the order posited by Markowitz. The only difference is that the decision utility is a function of outcomes expressed in relative rather than absolute terms as assumed by Markowitz.

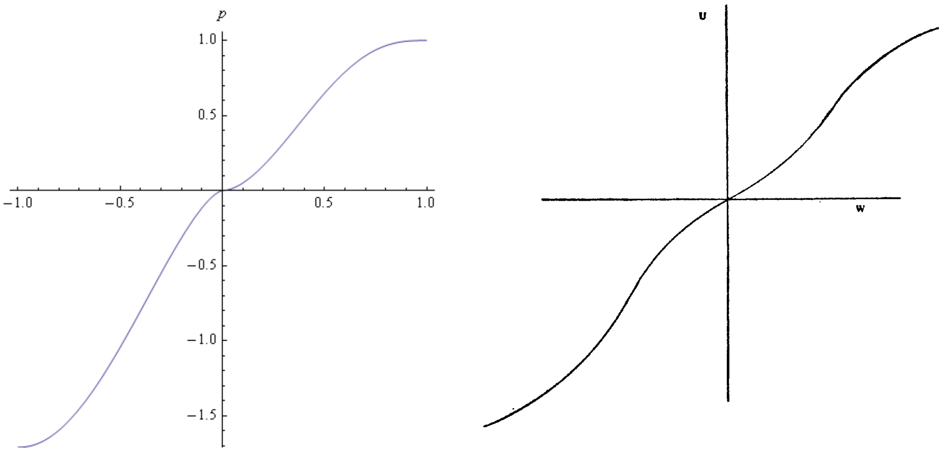

Figure 6.1 Left - Decision utility function $p = D(r)$ for gains and losses; Right - Utility function as presented in the Markowitz hypothesis (1952).

The utility shape presented in Figure 6.1 differs substantially from the one postulated by Prospect Theory, which is concave for gains and convex for losses. Nevertheless, it does offer an alternative explanation of the behaviors described by this theory. Cumulative Prospect Theory states that, when it comes to gains, people are risk seeking in case of low probabilities, and risk averse in case of high probabilities. Conversely, so far as losses are concerned, people are said to be risk seeking in case of high probabilities, and risk averse in case of low probabilities. This pattern, known as “the fourfold pattern of risk attitude”, is described by a combination of the value and probability weighting functions. The pattern finds, however, a simpler explanation in the use of the decision utility curve. This curve consists of four parts - convex, concave, convex, and concave - that correspond exactly with the "fourfold pattern".

Markowitz's assumption that the utility function is defined for wealth changes expressed in monetary terms precluded his curve (however tempting its shape) from being able to explain the paradoxes observed in lottery experiments. This is what led Kahneman and Tversky to reject the Markowitz hypothesis and develop Prospect Theory. Such paradoxes can, however, be explained using the Markowitz curve once the outcomes are framed, i.e. expressed relatively. This has already been shown for the Common Ratio Effect and will be demonstrated again in Point 11 when the Allais Paradox is explained.

7 Decision Utility as the Underlying Theory

Let us define the relative outcome r for a lottery with non-negative outcomes in a more general way than (4.1):

$$r = \frac{v(ce) - v(P_{min})}{v(P_{max}) - v(P_{min})} \tag{7.1}$$

where v denotes perception utility. The function v may be understood as the expression of how given monetary amounts are perceived by a subject. It has to be reminded that linearly perceived outcomes were assumed when deriving decision utility in Point 4. Although this approach works well in most cases, a nonlinear function, especially for wider outcome ranges, may generally be assumed. This does not change the essence of the decision utility model.

Substituting (7.1) into (4.2) results in:

$$p = D\left[\frac{v(ce) - v(P_{min})}{v(P_{max}) - v(P_{min})}\right] \tag{7.2}$$

which can alternatively be presented as:

$$D^{-1}(p) = \frac{v(ce) - v(P_{min})}{v(P_{max}) - v(P_{min})} \tag{7.3}$$

A simple rearrangement of (7.3) leads to:

$$v(ce) = v(P_{min})\left[1 - D^{-1}(p)\right] + v(P_{max})D^{-1}(p) \tag{7.4}$$

Let us assume that v denotes the Prospect Theory value function and D^{-1} denotes its probability weighting function w. Thus (7.4) states that the certainty equivalent value equals the two-outcome lottery value, according to Prospect Theory. It turns out that the Prospect Theory model can easily be derived from Decision Utility Theory.

The correspondence between theories discussed can now be summarized. The basic Expected Utility Theory model for a two-outcome lottery[7] can be presented as follows:

$$p = r \tag{7.5}$$

This results directly from the von Neumann-Morgenstern method (2.1) and the relative outcome definition (7.1). The basic Decision Utility Theory model can be presented as:

$$p = D(r) \tag{7.6}$$

where D denotes decision utility. This demonstrates the difference: probability p is defined by function $D(r)$ rather than by r, as assumed by von Neumann and Morgenstern, and shows that Expected Utility is a special case of Decision Utility Theory. The Prospect Theory model can be presented using (7.3) as:

$$w(p) = r \tag{7.7}$$

where w denotes the probability weighting function. This looks only like the inversion of (7.6) as (7.7) leads to different model properties. This important subject will be discussed in detail in the next Points.

Let us now describe the correspondence between decision and perception utility, and let us try to justify their names. Following Birnbaum (1974), the function ordinarily observed in a psychophysical experiment is the composition

$$\mathrm{R} = \mathrm{J}\left[\mathrm{H}(\phi)\right] \tag{7.8}$$

where the response function $\mathrm{R} = \mathrm{J}(\Psi)$ relates the responses R to the impressions Ψ, and the psychophysical function $\Psi = \mathrm{H}(\phi)$ relates the psychological impressions Ψ to the physical measures of the stimuli ϕ. The decision utility model (7.2) has exactly the same form as (7.8), with probability p as the response R, decision utility D as the response function J, perception utility v as the psychophysical function H, and their respective outcomes as the physical measures of the stimuli ϕ. The last of these is obviously money.

[7] Please note that (2.1) can be rearranged to: $u(A) = (1-p)u(B) + p\,u(C)$ which can be derived directly from the expected utility formula (1.1).

"Perception utility" is used here instead of "psychophysical function", as the von Neumann-Morgenstern utility u and the Prospect Theory value function v seem to express more complex phenomena than psychophysical impressions only. This, however, suggests yet another layer in the (7.8) model. The decision framing described by (7.1) or (4.1) can be regarded as the next layer in this multi-layer decision making model.

8 Stochastic Dominance Violations

The summary presented in Point 7 may lead to the conclusion that Decision Utility and Prospect Theory are alternative models, the only difference being the inverse representation (cf (7.6) and (7.7)). This conclusion is only partially legitimate. The representations (7.3) and (7.4) break with the axioms of classical probability, as a probability function appears there. This introduces stochastic dominance violations which could have only been solved by Cumulative Prospect Theory. In consequence, this theory departs from the simple decision utility model. The considerations presented further in this paper are crucial to understanding this fundamental difference.

Let us consider a lottery with an outcome of \$100 having a probability of 0.5. This case could be presented as an urn with two balls: one winning and one losing. The lottery value can be expressed as:

$$u_1(y) = w(0.5)\,v(100) \approx 0.4\,v(100) \tag{8.1}$$

Let us now consider an outcome of \$100 broken down into 50 outcomes of \$100 with probability of 0.01 each. This case could be presented as an urn with 100 balls: 50 winning and 50 losing, which is essentially the same lottery as the first one. Small probabilities are much overweighted according to Prospect Theory, so the fifty outcomes calculated separately give a greater lottery value than does the single outcome:

$$u_2(y) = 50\,w(0.01)\,v(100) \approx 50 \cdot 0.05\,v(100) \approx 2.5\,v(100) \tag{8.2}$$

This leads to the nonsensical conclusion that a lottery with \$100 split into 50 outcomes is not only more than 6 times better than a lottery with a single outcome, but also 2.5 times better than a certain \$100. Such paradoxes where one lottery is valued more than another, even though it is clearly inferior, are referred to as stochastic dominance violations. The origin of the problem is clear: although the probabilities sum to 1, their weights do not. This makes the value of a lottery dependent on how it is represented.

Many other examples of stochastic dominance violations can be given (Weber, 1994). At first, Prospect Theory assumed that stochastic dominance problems could be avoided with the assistance of mental operations carried out during the editing phase of decision making, i.e. before the lottery valuation proper. These operations detect and eliminate violations; for instance the calculation (8.2) is prohibited. The solution was only partially satisfactory as the theory remained restricted to two-outcome lotteries (a popular Lotto cannot be described by Prospect Theory). Therefore, the early probability weighting concept was replaced by applying the probability weighting to cumulative probabilities rather than individual ones (Quiggin, 1982). Rank-Dependent Expected Utility Theory, which introduced this solution, requires, however, that the weighting is no longer solely a function of probability, but of outcome ranking as well. The theory assumes that the utility of a prospect $X=(p_1,x_1;...;p_n,x_n)$, with outcomes ordered in increasing order $x_1<...<x_n$, is defined as:

$$RDU(X)=\sum_{i=1}^{n}u(x_i)w(p_i,X) \tag{8.3}$$

where

$$w(p_i,X)=w(p_i+...+p_n)-w(p_{i+1}+...+p_n) \tag{8.4}$$

Further explanation is cited from Weber (1994): "*The first expression in (8.4) is the sum over probabilities of all outcomes that are at least as great as x_i; the second expression is the sum over the probabilities of all outcomes that are greater than x_i. The dependence on the rank of x_i comes about because different probability values enter into the two summations, depending on the rank of x_i. As the result the probability weight $w(p_i, X)$ given to probability p_i depends critically on the rank of the associated outcome x_i*".

The Rank-Dependent Expected Utility model has been further implemented virtually unchanged by Cumulative Prospect Theory (1992). The considerations presented in this Point show that the introduction of probability weighting to description models, initially regarded as the simplest explanation of the behaviors observed in lottery experiments, appears to be an "evolutionary" trap, which has led those models to complex and unwieldy forms. Per contra, the decision utility model preserves its simple form even in the case of multi-outcome lotteries. This is because stochastic dominance violations are not introduced by decision utility. This will be presented in the next Point.

9 Expected Decision Utility

Decision utility considers probabilities within the framework of the classical axioms (cf (7.6)). This allows the general formula for multi-outcome lotteries to be derived in a straightforward way. It is assumed that for each multi-outcome lottery there exists an equivalent lottery with two outcomes of 0 and 1. Each relative outcome r_i of the multi-outcome lottery is substituted by a lottery whose outcomes are 0 and 1 and whose probability of winning is $p_i^{'} = D(r_i)$. The joint probability of winning the equivalent lottery is therefore

$$p_{eq} = \sum_{i=1}^{n} p_i^{'} p_i \tag{9.1}$$

where p_i denotes the probability of the outcome r_i in the multi-outcome lottery. This leads to:

$$p_{eq} = \sum_{i=1}^{n} D(r_i) p_i \tag{9.2}$$

This ends the derivation as the equivalent probability p_{eq} expresses the decision utility of the multi-outcome lottery. Please note that the obtained expected decision utility formula is almost identical with (1.1) postulated by Expected Utility Theory. The difference is that decision, rather than perception utility is used, and the considered outcomes are expressed in relative rather than in absolute terms. Please also note that in the case of two-outcome lotteries, (9.2) always reduces to $p_{eq} = p$, where p is the probability of winning the greater outcome.

Let us now consider a relative outcome of r_i having a probability of p_i. Its individual contribution to (9.2) is given by:

$$p_{eq,i} = D(r_i) p_i \tag{9.3}$$

Let us now assume that this outcome is presented as k outcomes each having a probability of p_i / k. These would be represented within (9.2) as follows:

$$p_{eq,i}^{'} = D(r_i)\frac{p_i}{k} + ... + D(r_i)\frac{p_i}{k} = \sum_{j=1}^{k} D(r_i)\frac{p_i}{k} = D(r_i) p_i \sum_{j=1}^{k}\frac{1}{k} = D(r_i) p_i \tag{9.4}$$

which is exactly the same as (9.3). It follows that the way in which a lottery is represented has no impact on its equivalent probability p_{eq}. Decision Utility Theory does not, therefore, introduce stochastic dominance violations. As a result, decision utility presents a very different model from the one offered by Prospect Theory – especially its Cumulative version – and allows lotteries to be valued in a similar way as postulated by Expected Utility Theory. An editing phase, probability weighting, cumulative probability representation and outcome ranking are all redundant for this purpose.

10 Decision Utility vs. Cumulative Prospect Theory

Decision Utility and Prospect Theory are pretty much the same description models for two-outcome lotteries. The main difference is that the former does not require an editing phase to exclude stochastic dominance violations. The models substantially diverge for multi-outcome lotteries, in which case Cumulative Prospect Theory has to be used. This may raise the question of which model, decision utility or cumulative prospect, is correct in such applications. This could be checked using experimental data, and the existence of a wide literature on Cumulative Prospect Theory should easily validate this model. This, however, appears to be problematic. It is a frequently overlooked fact that the Cumulative Prospect Theory parameters were derived by only using two-outcome lotteries (Tversky, Kahneman, 1992). Applying the theory to multi-outcome lotteries has only been justified by a mathematical proof. Evidence of the Cumulative Prospect Theory model having been estimated using lotteries with more than two outcomes is conspicuously lacking[8].

In any case, using experimental data to check both models would have the disadvantage of being sensitive to the data collected. A theoretical consideration is therefore presented. Let us consider a lottery with two outcomes: \$0 and \$100, each with a 0.5 probability of occurrence. The lottery certainty equivalent was observed to be \$36 (Tversky, Kahneman, 1992). Let us next consider a second lottery having three outcomes: \$0, \$36 and \$100, where the middle outcome of \$36 is the same as the certainty equivalent of the first lottery. The outcomes have respective probabilities of $(1 - p)/2$, p, and $(1 - p)/2$. Logical reasoning leads to the conclusion that the second lottery certainty equivalent should always assume a value of \$36, whatever the value of probability p. This reasoning proceeds as follows. As p approaches 0, the second lottery reduces to the first lottery, whose certainty equivalent is \$36. As p approaches 1, the second lottery reduces to a sure payment of \$36. For any other probability p in the [0, 1] range, the second lottery may be regarded as a mixture of two prospects having the same equivalent of \$36. There is,

[8] The only trace of using multi-outcome lotteries has been found in Gonzales and Wu (1999), who stated (footnote 6): "*In addition to the two outcome gambles, 36 three-outcome gambles were included. Data from these gambles will be presented elsewhere*". We are not going to speculate why these results, which would confirm the applicability of Cumulative Prospect Theory to multi-outcome lotteries, have yet to be presented.

therefore, no reason to predict that its certainty equivalent will assume a value other than \$36. This is the precise value predicted by the decision utility model. Let us note that according to (9.2):

$$p_{eq} = D(0)\frac{1-p}{2} + D\left(\frac{36}{100}\right)p + D(1)\frac{1-p}{2} = 0 + \frac{1}{2}p + \frac{1-p}{2} = \frac{1}{2} \tag{10.1}$$

Thus, according to (4.5), the lottery certainty equivalent assumes a value of \$36 for any probability p. Let us now derive the solution using Cumulative Prospect Theory. The value of the second lottery is expressed as:

$$\begin{aligned} \mathrm{V}(y) = \mathrm{v}(100)\left[w\left((1-p)/2\right) - w(0)\right] + \mathrm{v}(36)\left[w\left((1-p)/2+p\right) - w\left((1-p)/2\right)\right] \\ + \mathrm{v}(0)\left[w(1) - w\left((1-p)/2+p\right)\right] \end{aligned} \tag{10.2}$$

This value should equal the certainty equivalent value $v(ce)$. Applying the functional forms proposed by Cumulative Prospect Theory leads, after simplification, to the following expression for the lottery certainty equivalent ce:

$$ce = \left\{ \frac{\left(100^{\alpha} - 36^{\alpha}\right)\left((1-p)/2\right)^{\gamma} + 36^{\alpha}\left((1+p)/2\right)^{\gamma}}{\left[\left((1-p)/2\right)^{\gamma} + \left((1+p)/2\right)^{\gamma}\right]^{1/\gamma}} \right\}^{\frac{1}{\alpha}} \tag{10.3}$$

where α and γ denote the parameters of the value and probability weighting functions. The certainty equivalent (10.3) can only be calculated numerically and is presented in Figure 10.1 for probability p in the [0, 1] range. The certainty equivalent obtained using the decision utility model is a constant \$36, as explained.

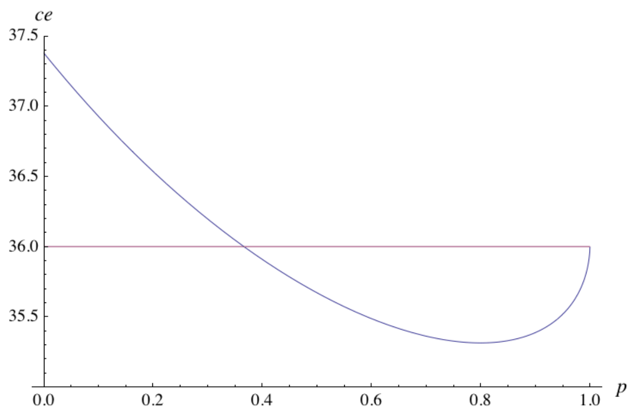

Figure 10.1 Certainty equivalent of the considered three-outcome lottery as a function of probability *p* of a \$36 outcome. The solution for the decision utility model is a constant \$36. The solution for CPT is marked in blue.

The solution obtained using Cumulative Prospect Theory depends on probability p and model parameters α and γ. It comes out that the solution of stochastic dominance problems has led to the model producing pretty "strange" predictions even in a simple multi-outcome lottery case. That this result is largely inexplicable calls the correctness and applicability of this theory into question.

11 Allais Paradox

Once the formula for the multi-outcome lottery calculation has been derived, the Allais paradox can now be explained. It is presented in the form proposed by Kahneman and Tversky (1979).

Problem 3: Choose between

A: 2,500 with a probability of 0.33	B: 2,400 with a probability of 1.00
2,400 with a probability of 0.66	
0 with a probability of 0.01	

Problem 4: Choose between

C: 2,500 with a probability of 0.33	D: 2,400 with a probability of 0.34
0 with a probability of 0.67	0 with a probability of 0.66

Experiments consistently show that most people choose option B in Problem 3 and option C in Problem 4, which contradicts the choices predicted by Expected Utility Theory. The solution using decision utility proceeds as follows. The certainty equivalents of options A and B are determined using (9.2) and (4.5):

$$ce_A = 2500\,D^{-1}\left[0.33D\left(\frac{2500}{2500}\right)+0.66D\left(\frac{2400}{2500}\right)\right] \qquad (11.1)$$

$$ce_B = 2400\,D^{-1}\left(1\right) \qquad (11.2)$$

Solving the inequality $ce_A < ce_B$ results in the condition:

$$D\left(\frac{2400}{2500}\right) > \frac{0.33}{0.34} \quad \text{or} \quad D\left(0.96\right) > 0.97 \qquad (11.3)$$

As 0.97 is greater than 0.96, a risk aversion attitude for relative outcomes approaching a value of 1 is required to solve the problem (for more detail, refer Point 13). As:

$$ce_C = 2500\,D^{-1}\left(0.33\right) \qquad (11.4)$$

$$ce_D = 2400\,D^{-1}(0.34) \tag{11.5}$$

the second inequality $ce_C > ce_D$, i.e.:

$$D^{-1}(0.33) > 0.96\,D^{-1}(0.34) \tag{11.6}$$

presents another condition for decision utility. Both conditions are fulfilled by the decision utility determined in Point 4. More general analysis (not presented here) shows that only an S-shaped decision utility curve can explain both the Allais Paradox and the Common Ratio Effect.

12 Decision Utility and Choice Heuristics

Decision Utility Theory offers possibly a more comprehensive explanation of decision making than is the case of Prospect Theory. Applying the rule of utility maximization to the decision utility, which is expressed in probability terms, leads to the conclusion that the option with the greatest probability should always be chosen. This conclusion has already been stated when presenting the von Neumann-Morgenstern method in Point 2. Such way of proceeding seems to be wrong when the outcome ranges differ. The objection can be answered that the lottery valuation consists of two stages, according to Decision Utility Theory (refer Point 4). The first stage is decision utility maximization, and the second is rescaling the result to absolute values (4.4). Lotteries can thus be compared by their certainty equivalents.

The two-stage valuation process predicts, however, that people might not weigh up prospective outcomes against the probabilities of their occurrence, contrary to the expected utility paradigm. Maximizing decision utility indicates that a safer option may be preferred, even when the outcome ranges differ. Some proverbs can be cited here: a bird in the hand is worth two in the bush; better an egg today than a hen tomorrow. On the other hand, for similar (less known, unknown) probabilities, the maximum outcome value proves to be decisive in deciding which choice to make. An interesting case is that many people buy lottery tickets when there is an accumulated prize of 10 million dollars, but not when the prize is “only” 1 million.

These conclusions would be consistent with Brandstätter et al. (2006), who postulated that people use simple heuristics rather than the expected utility calculation when making choice decisions. These heuristics may be presented in a simple manner: (1) choose the safe option whose worst outcome exceeds the aspiration level; (2) if no such option exists, then choose the one that has the greatest probability of winning; (3) if the probabilities are similar, choose the option with the greatest maximum outcome. After analyzing many experimental data, Brandstätter et al. concluded that these simple heuristics predict choices much better than theories employing the expected utility paradigm (these include Prospect Theory). Very clearly, probability appears to be the main factor affecting choice decisions in many situations. In many others, the maximum outcome decides.

The two-stage valuation presented by Decision Utility Theory may thus serve to explain preferences for probabilities or outcomes in the decision making process. Decision heuristics obtained this way may create yet another layer in the multi-layer decision making process presented in Point 7.

13 Risk Attitudes – Back to Basics

Decision utility differs from the Expected Utility and Prospect Theory approaches yet in the way it defines attitude to risk. In the classical economic approach, this attitude is defined by the shape of the utility curve. Hence the risk measures proposed by Arrow and Pratt. However, in the case of decision utility, attitude to risk is defined by a single point; more precisely by its placement relative to the straight line $p = r$, referred to here as the Expected Value Line (EVL)[9]. If a given point $\{r, p\}$ lies above the EVL, then the individual is risk averse, as the certainty equivalent r is lower than the lottery expected value (see Figure 13.1).

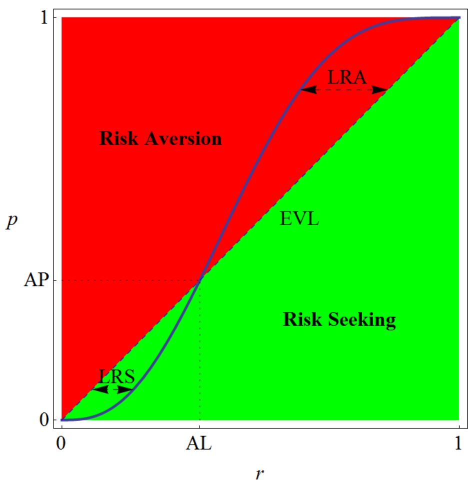

Figure 13.1 Risk Seeking and Risk Aversion areas plotted together with the Expected Value Line EVL, Aspiration Level AL, Aspiration Probability AP, Local Risk Aversion, and Local Risk Seeking measures.

If this point is below the EVL then the individual is risk seeking as the certainty equiva-

[9] A linear perception function v is assumed in the presented considerations.

lent is greater than the lottery expected value. Finally, if the point lies on the EVL then the individual is risk neutral as the certainty equivalent and the lottery expected value are equal. As the location of a single point is sufficient to determine risk attitude, the shape of the decision utility curve is of no importance. Quite surprisingly, a *convex* section of the decision utility curve can be imagined above the EVL to describe a very specific type of risk *aversion* attitude.In this context, the point where the decision utility curve intersects the EVL assumes great importance. This point can be called the Aspiration Level (AL), because this is where people change their attitude to risk from one of risk seeking to one of risk aversion. Participating in lotteries is preferred for relative outcomes less than AL. Conversely, a sure payment is preferred if it is greater than AL. Similarly, the Aspiration Probability (AP) can be defined, which assumes the same numerical value as AL, but expresses probabilities rather than relative outcomes. The Local Risk Aversion (LRA) measure can be then defined as the difference between the point $\{p, p\}$ located on EVL and the given point $\{r, p\}$. The Local Risk Seeking (LRS) measure is defined with the opposite sign.

General risk attitude measures can be defined as well. Assuming a set of points $\{r_i, p_i\}$ distributed symmetrically around a probability of 0.5, let us consider the average value of the relative certainty equivalents:

$$RS = \frac{1}{n}\sum_{k=1}^{n} r_k \tag{12.1}$$

Clearly, the higher the relative certainty equivalents, the more risk seeking the attitude. This measure therefore reflects the general risk-seeking attitude of a given subject. *RS* assumes a value of 0 for the maximum risk aversion attitude, i.e. all where all the certainty equivalents are equal to the minimum lottery outcomes, and a value of 1 for the maximum risk seeking attitude, i.e. where all the certainty equivalents are equal to the maximum lottery outcomes. *RS* assumes a value of 0.5 for general risk neutrality. The symmetry of the assumed values allows the general Risk Aversion *RA* measure to be defined as:

$$RA = 1 - RS \tag{12.2}$$

A subject is therefore said to be risk averse when the relative certainty equivalents are on average less than 0.5 and to be risk seeking when the relative certainty equivalents are on average greater than 0.5. Let us next consider a smooth distribution of $\{r_i, p_i\}$ points. *RS* can be presented as:

$$RS = \int_0^1 r\,dp = \int_0^1 D^{-1}(p)\,dp \tag{12.3}$$

which appears to be the area to the left of the decision utility function (see Figure 13.2). Clearly, this area can also be seen as the area *above* the decision utility.

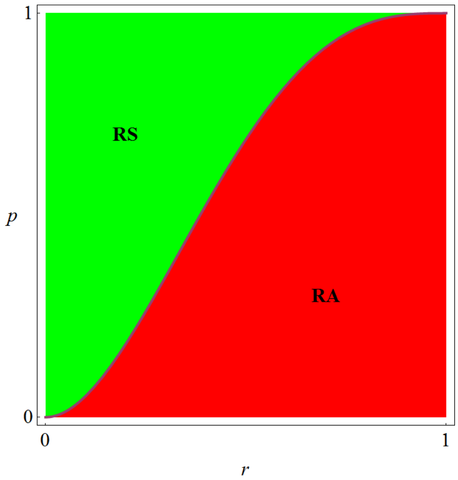

Figure 13.2 Risk Aversion RA measure defined as the area below the decision utility curve; Risk Seeking RS measure defined as the area above the decision utility curve. Both areas sum to 1.

As the area of the square is 1, general risk aversion *RA* can be presented as the area below the decision utility function:

$$RA = 1 - RS = \int_0^1 D(r)\,dr \tag{12.4}$$

It follows that the further left the decision utility function is shifted, the higher the *RA* measure and, conversely, the further right, the higher the *RS* measure[10].

14 Probability Weighting Puzzle

Prospect Theory has always been thought of as a psychological theory. The introduction of a reference point, the distinction between gains and losses, and the concept of loss aversion all support this common view[11]. Probability weighting is likewise considered a psychological concept. There is, however, a great deal of misunderstanding regarding this subject. To quote Kah-

[10] Calculating risk attitude measures is extremely easy once the Cumulative Beta Distribution is used to describe the decision utility function (as in Point 4). In this case $RA = \frac{\beta}{\alpha + \beta}$ and $RS = \frac{\alpha}{\alpha + \beta}$, where α and β denote the function parameters.

[11] In fact these were all first introduced by Markowitz in 1952.

neman and Tversky (1979): "*It is important to* *distinguish* ***overweighting****, which refers to a property of decision weights, from* ***overestimation*** *that is commonly found in the assessment of the probability of rare events*" (emphases added). The authors of Prospect Theory are thus fully aware that these are two different phenomena and in no way claim that overweighting has any psychological basis. As overestimation "*does not arise in the context, where the subject is assumed to adopt the stated value of probability*", they explain lottery results in terms of overweighting. However, the nature of overweighting remains unspecified.

Despite this very clear statement, the all too common failure to distinguish between overweighting (a mathematical concept to explain the results of lottery experiments regarding events whose probabilities are known) and overestimation (which can be referred to as a kind of subjective view of events whose probabilities are not known) leads to the generally accepted view that probability weighting has a profound psychological justification; moreover, both phenomena are thought to result from the same underlying psychological process and to only manifest themselves in different situations. Despite the extensive research that has been conducted since Prospect Theory was introduced, the basic question, however, remains unanswered: *Why and how do people overweight small probabilities and underweight high probabilities when the probabilities are known?* Until an answer to this question is given, and a psychological explanation of this phenomenon presented, probability weighting cannot be regarded as the psychological explanation of the decision making behaviors observed in lottery experiments.

This paper shows that decision utility can explain these behaviors. For instance, the "fourfold pattern of risk attitudes" can be explained in several ways. The first was presented in Point 6 using the double S-shape of the decision utility curve. Another explanation deploys the concept of aspiration level introduced in Point 13: people take risks until prospective gains are lower than their aspiration levels; people will avoid risks until prospective gains exceed their aspiration levels. This explanation is not new and can be found in the Simon's theory of bounded rationality (Simon, 1982). The first heuristic rule of Brandstätter et al. should also be reminded: choose the lottery if its worst outcome exceeds the aspiration level. Thus, the explanation presented has a very solid psychological justification as aspiration level is a well described concept in the psychology and economics literature. This concept easily explains why people start new businesses, play the stock market and change their partners. It also explains why they sell businesses, buy bonds and stop looking for a new job. Decision Utility Theory is, however, much more precise than simple heuristics, because it can numerically measure the behaviors described.

A similar explanation can be given for losses, although "desperation", to use the anti-aspiration term, might be regarded as too strong[12]. However that may be, people will avoid risks while prospective losses do not exceed their desperation level, but will take risks once this level is exceeded. This explains why people buy insurance, but some people steal food when they have no money to buy it. Please note, that the presented explanations assume that the aspiration and desperation levels are not constant and may vary depending on the context. This should be obvious as decisions are always framed according to Decision Utility Theory.

As shown, a wide range of behaviors can be explained in this very psychological way. The same behaviors, when observed from the Prospect Theory point of view, would have to be interpreted by claiming that *people behave as if they were distorting probabilities* when making decisions (even if those probabilities are known). Which explanation is the more plausible is left for the reader to decide.

15 Summary

By making use of well known experimental data, this paper has demonstrated the origin of the problems of explaining the behaviors observed in lottery experiments. This is the assumption of a single utility being able to describe them. A single utility defined for outcomes expressed in absolute terms makes, however, lottery paradoxes impossible to explain. This may raise doubts as to whether the observed behaviors are really paradoxical or merely interpreted as such when a single utility is applied to describe them.

A single utility is the core assumption underlying both Expected Utility Theory and Prospect Theory. The relative approach introduced by the latter only partially helps to explain lottery results. This is because gains and losses, although considered relative to a reference point, are still expressed in terms of monetary amounts. Behavioral paradoxes have only ever been explained by recourse to the probability weighting concept. Being divorced from the axioms of classical probability, the Prospect Theory approach has led to stochastic dominance violations. These can only be avoided by adding an Editing phase. As this solution has proved inadequate, the cumulative probability lottery representation has been additionally applied. This requires the probability weights to be dependent on outcome rankings and not just probabilities. What's more,

[12] The term "desperation" appeared in an unpublished work of Markowitz on the taxonomy of utility shapes.

the solution finally obtained to eliminate stochastic dominance problems, produces largely inexplicable predictions. Building the theory around a single utility thus led to a series of problems, which have had to be resolved using more and more complex methods.

All the presented problems can be avoided by using decision utility, which is defined for framed, i.e. relatively expressed outcomes. One big advantage of this approach is its simplicity. It is a single-equation theory that preserves the axioms of classical probability. Stochastic dominance violations are, therefore, not introduced, and neither an editing phase, nor probability weighting, nor cumulative probability representation, nor outcome ranking are needed to describe lottery experiments. Lotteries can be valued using the expected decision utility formula, similar to that proposed by von Neumann and Morgenstern. The shape of the decision utility curve, which resembles the one hypothesized by Markowitz (1952), explains major behavioral paradoxes. In fact, such behaviors are not regarded as paradoxical by Decision Utility Theory, as they can be easily predicted. We shall call this state of affairs, in which some sets of rationality axioms do consider them as paradoxes, the *irrationality illusion*.

It needs to be stressed that Decision Utility Theory has not been derived to explain any new psychological or economic phenomena. Although the theory predicts that people might not consider probabilities and outcomes in a joint manner, its essence and its main contribution lie in a new way of describing behavior. This radical change may be hyperbolically compared to the Copernican Revolution.

Expected Utility Theory assumes that $p = r$, where r is defined by (7.1). This model, however, is too simple to explain lottery experiments for all probability values, because it does not include decision utility, as defined in this paper.

Prospect Theory and its Cumulative version assume that $w(p) = r$, where w denotes the probability weighting function. This model leads to many description complexities, in much the same way that "geo-centrism" leads to having the planets travel in deferents and epicycles.

Decision Utility Theory assumes that $p = D(r)$, where D denotes the decision utility function of the shape predicted by Markowitz. This model preserves the original von Neumann-Morgenstern concept of utility expressed as a probability and puts everything in its rightful place. In doing so, it offers a simple description of lottery behaviors, just as the heliocentric system offers a simple description of planetary movements.

References:

Allais, M., (1953). *Le comportement de l'homme rationnel devant le risque: critique des postulats etaxiomes de l'école Américaine*. Econometrica 21, pp 503-546.

Birnbaum, M. H., (1974). *Using contextual effects to derive psychophysical scales.* Perception & Psychophysics, Vol. 15, No. 1 pp. 89-96.

Brandstätter, E., Gigerenzer, G., Hertwig, R., (2006). *The Priority Heuristic: Making Choices Without Trade-Offs*. Psychological Review, 113(2), pp. 409-432.

Gonzales, R., Wu, G., (1999). *On the Shape of the Probability Weighting Function.* Cognitive Psychology, 38, pp 129-166.

Kahneman, D., Tversky, A., (1979). *Prospect theory: An analysis of decisions under risk.* Econometrica, 47, pp 313-327.

Kahneman, D., Wakker, P., Sarin, R., (1997). *Back to Bentham? Explorations of Experienced Utility*. The Quarterly Journal of Economics, May, pp 375-405.

Kahneman, D., (1999). *Objective Happiness*. In Kahneman, D., Diener, E., Schwartz, N. (editors) *Well-Being. The Foundation of Hedonic Psychology*. Russell Sage Foundation, pp 3-25.

Kahneman, D., Thaler, R., (2006). *Anomalies: Utility Maximization and Experienced Utility*. Journal of Economic Perspective, vol 20, No.1, pp 221-234.

Markowitz H., (1952). *The Utility of Wealth.* Journal of Political Economy, Vol. 60, pp. 151-158.

Read, D., (2004). *Utility theory from Jeremy Bentham to Daniel Kahneman*. Working Paper LSEOR 04-64, London School of Economics and Political Science.

Simon, H., (1982). *Theories of bounded rationality.* In H. Simon (editor), *Models of bounded rationality. Behavioral economics and business organization*. Cambridge, MA, MIT Press. Vol. 2, pp 408-423.

Tversky A., Kahneman D., (1992). *Advances in Prospect Theory: Cumulative Representation of Uncertainty*. Journal of Risk and Uncertainty, vol. 5(4), October, pp 297-323.

Van Praag, B. M. S., Frijters, P., (1999). *The Measurement of Welfare and Well-Being: The Leyden Approach*. In Kahneman, D., Diener, E., Schwartz N. (editors), *Well-Being. The Foundation of Hedonic Psychology*. Russell Sage Foundation, pp 413-433.

Von Neumann J., Morgenstern O., (1944). *The Theory of Games and Economic Behavior*, Princeton University Press.

Weber, E., (1994). *From Subjective Probabilities to Decision Weights: The Effect of Asymmetric Loss Functions on the Evaluation of Uncertain Outcomes and Events.* Psychological Bulletin, Vol. 115, No 2, pp 228-242.