

TRANSFORMING PROBABILITIES WITHOUT VIOLATING STOCHASTIC DOMINANCE

*Peter Wakker**

University of Nijmegen

*Nijmegen Institute for Cognition research and
Information technology (NICI)*

Abstract

The idea of expected utility, to transform payments into their utilities before calculating expectation, traces back at least to Bernoulli (1738). It is a very natural idea to transform, analogously, probabilities. This paper gives heuristic visual arguments to show that the, at first sight, natural way to do this, at second thought seems questionable. At second thought a sound and natural way is the way of anticipated utility, as indicated by Quiggin (1982).

SECTION 1. INTRODUCTION

In this paper we present a new argument, heuristic and visual, for a new approach to decision making under risk, the 'anticipated utility' approach. It has been initiated by Quiggin (1982) and Yaari (1987a) (and Schmeidler, 1982), and deals with the idea to transform probabilities. While being very old, this idea until recently did not lead to sound theories. As we shall argue this is because, loosely spoken, the old approaches to transformed probabilities 'transformed the wrong probabilities' when calculating expectation. This prevented the idea from becoming full-blown. Only recently, through the references mentioned above, it has become understood how to integrate with respect to transformed probabilities, i.e., how to calculate expectations with respect to transformed probabilities. This is best done by means of the 'Choquet-integral', introduced by Choquet (1953-4), or its analogues as used in this paper. The purpose of this paper is to give a visual presentation by means of which the reader may come to the intuition that the anticipated utility approach, rather than older approaches, is a natural analogue of expected utility. Hopefully this will interest the reader in getting to know more about anticipated utility to decision making under risk, and to consult the references given at the end. Finally a remark on terminology. Instead of 'anticipated

*The research has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences. Author's address: University of Nijmegen (NICI), Department of Mathematical Psychology, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands

utility' one also finds in the literature terms such as 'rank-dependent expected utility', 'rank-ordered expected utility', 'expected utility with rank-dependent probabilities'. The latter term reflects best the idea of the new approach, but is not concise.

SECTION 2. LOTTERIES AND EXPECTATIONS

The simplest approach to decision making under risk consists of the maximization of 'expected value'. Let $(p_1; x_1, \dots, p_n; x_n)$ be a **lottery** yielding $\$x_j$ with probability p_j , $j = 1, \dots, n$. Then the **expected value** of the lottery is

$$\sum_{j=1}^n p_j x_j . \quad (1)$$

With the reader's indulgence we present, as a necessary preparation for the sequel, two elementary illustrations. In Figure 2.1a the shaded area indicates the expected value. The first column has height x_1 and breadth p_1 , hence its area gives the first term $p_1 x_1$ of the summation of expected value; etc. **Let us emphasize that we assume in the sequel that the outcomes x_1, \dots, x_n have been ordered so that $x_1 > x_2 > \dots > x_n$.**¹ Also we assume that all x_j 's are positive; this is merely done for convenience. Figure 2.1b shows an alternative illustration of expected value. It has simply been obtained from Figure 2.1a by rotating to the left and flipping (as the reader may imitate with this paper). Obviously, the size of the shaded area is not affected by these operations. If one considers the shape in Figure 2.1b as the graph of a function, denoted by G , then the area in it is the integral of G . This function G in fact is the 'decumulative distribution function' of the lottery; the **decumulative distribution function**² assigns to any real number t the probability that the lottery will yield at least $\$t$, i.e.,

$$G : t \mapsto P(\text{lottery yields } \$t \text{ or more}) .$$

Indeed, for every $t \in \mathfrak{R}_+$ the height of the graph is the sum of the probabilities of the x_j 's to the right of t . Hence the expected value of the lottery can also be obtained as the integral of the decumulative distribution function, i.e.,

$$\sum_{j=1}^n p_j x_j = \int_{\mathfrak{R}_+} G(t) dt .$$

(By elementary means, such as integration by parts, the above integral can be reduced to the more common $\int_{\mathfrak{R}_+} t f(t) dt$, with f the usual density function, so f the derivative of $-G$.)

A challenge to the maximization of expected value was posed by the 'St. Petersburg paradox'. Suppose $p_j = 2^{-j}$ for all j , and $x_j = 2^j$ for all j . Then the expected value of the lottery $(p_1; x_1, \dots, p_n; x_n, 2^{-n}; 0)$ is n . It turned out however that for large n

¹For formula (3) the inequalities $x_1 \neq x_2 \neq \dots \neq x_n$ will be essential. For section 4 the weak inequalities $x_1 \geq x_2 \geq \dots \geq x_n$ will be essential.

²Probably better known than the decumulative distribution function is the usual '(cumulative) distribution function', assigning $P(\text{lottery yields } \$t \text{ or less})$ to every real number t . For the visual presentation of this paper the decumulative distribution serves best; we are following Yaari (1987a) in this. Yaari also uses such a visual presentation to consider anticipated utility as dual to expected utility.

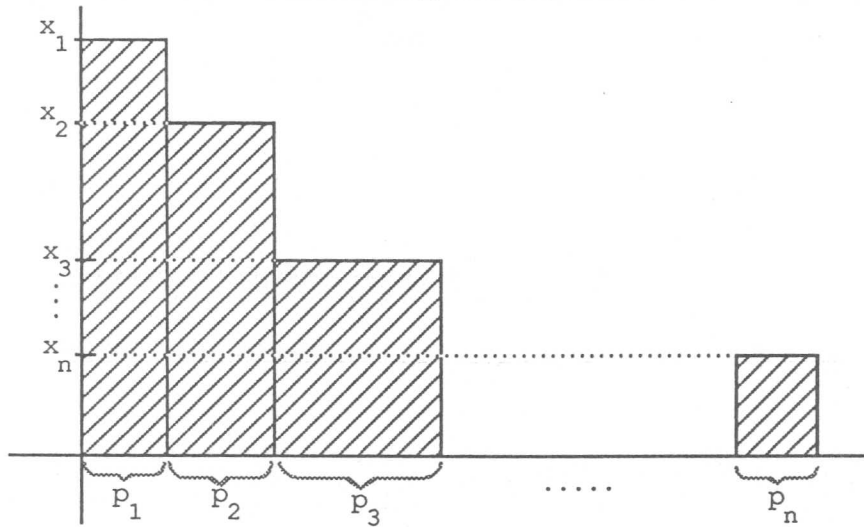



FIGURE 2.1a. *Expected value.* The area shaded by  gives the expected value.

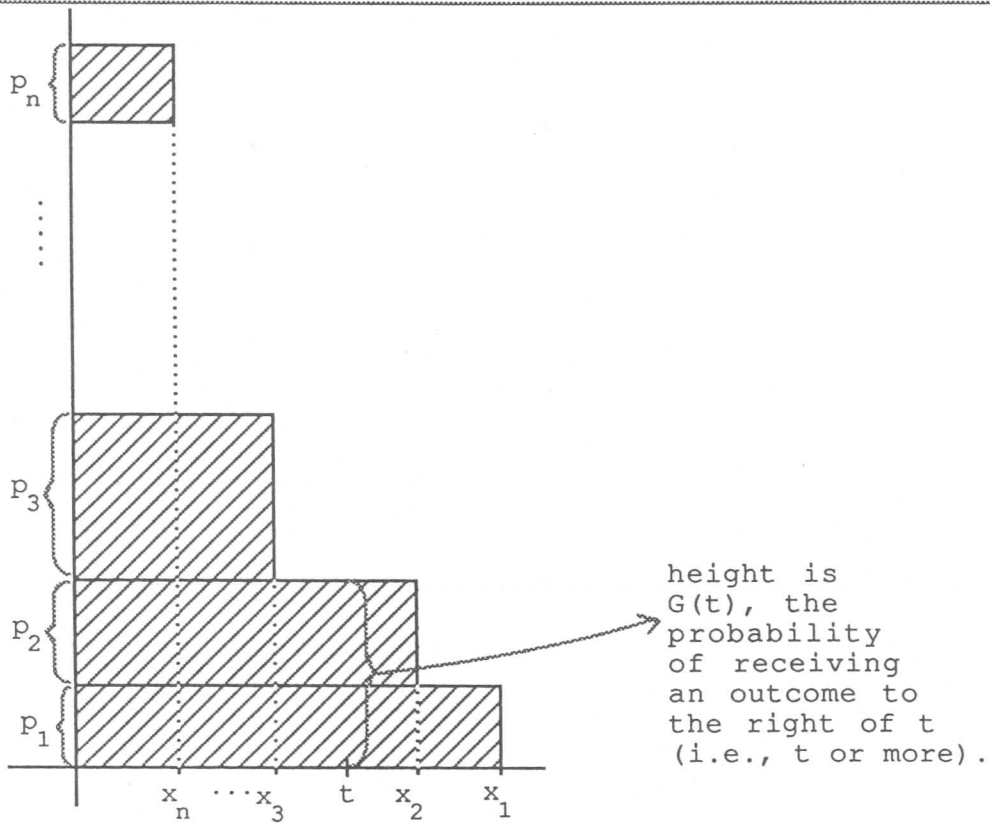



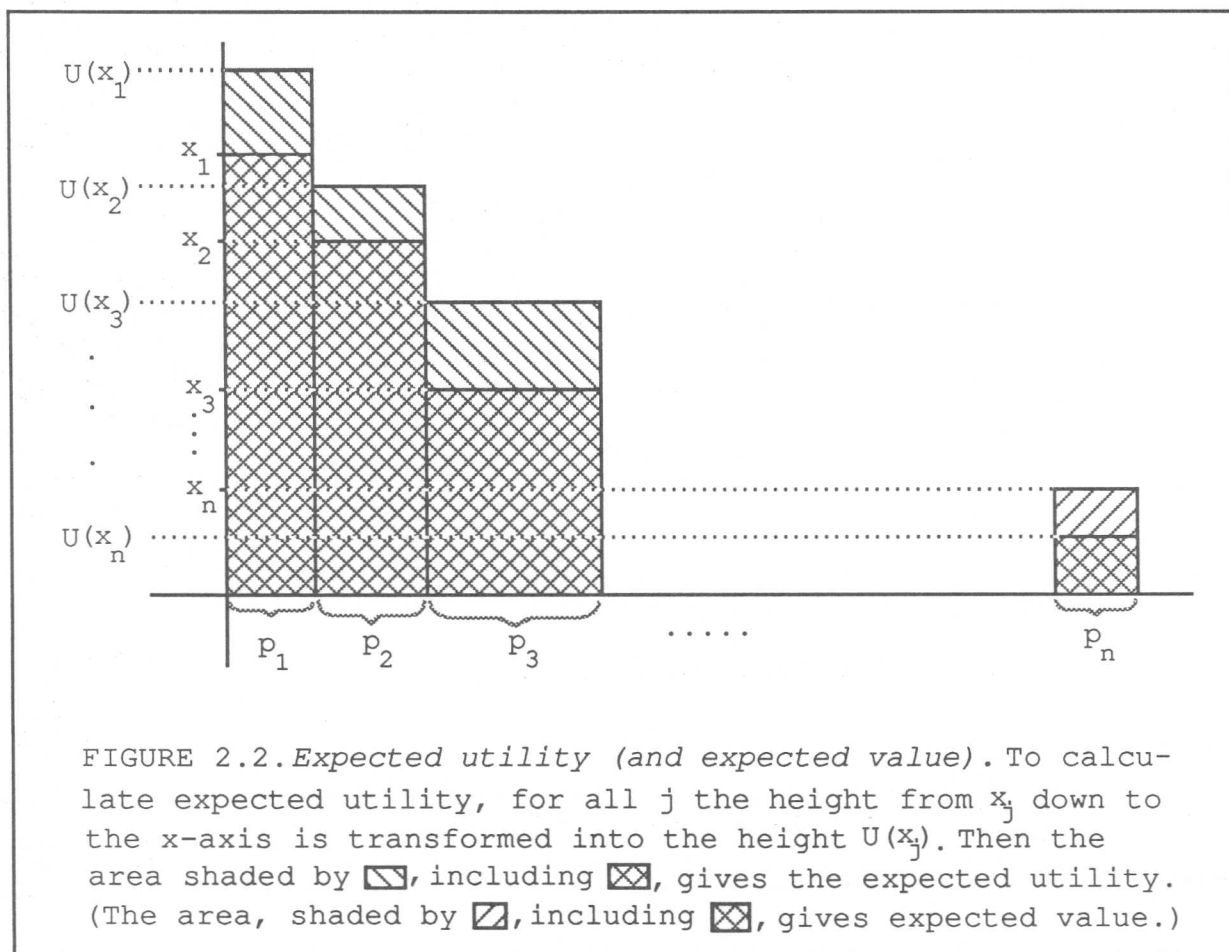
FIGURE 2.1b. *Alternative presentation of expected value.* Figure 2.1a has been rotated to the left, and flipped horizontally. Of course the area shaded by  still gives the expected value. The figure gives the graph of the decumulative distribution function G ; the area is its integral.

FIGURE 2.1. *Expected value represented as area.*

individuals were in no way willing to pay \$ n to participate in such a lottery. Usually not more than \$4 would be paid. Bernoulli's idea was that one should not calculate the expectation of the outcomes of money, but one should first transform the amounts of money by a 'utility function' U , and only then calculate expectation. As utility function Bernoulli proposed the logarithm. Indeed, $\sum_{j=1}^n 2^{-j} \log(2^j)$ tends to $\log(4)$ as n increases, so Bernoulli's approach indeed explains that people will not pay more than \$4 for participation in the lottery. (We prefer the presentation as above, with 'large' n , to the presentation with $n = \infty$. Intuitions of people are frequently misled by meaningless mathematical conventions concerning ∞ .)



With further indulgence of the reader, Figure 2.2 gives an elementary presentation of expected utility, again as a preparation for the intuitive argument to be developed. As compared to Figure 2.1a, for expected utility first the heights of the columns are transformed; the height x_1 is transformed into $U(x_1)$, the height x_2 (down to the x -axis!) is transformed into $U(x_2)$, etc. Finally, the **expected utility** of the lottery is the area shaded by $\\$ in Figure 2.2 (including the x -area), i.e.,

$$\sum_{j=1}^n p_j U(x_j) . \tag{2}$$

The expected utility approach became full-blown when it received behavioural foundations by von Neumann & Morgenstern (1944) and Savage (1954). It stirred statistical literature, a topic we shall not pursue here. Expected utility was used to explain phenomena related to risk. For instance 'risk aversion' was explained through concavity

of the utility function. If the utility function is concave then indeed a lottery is never preferred to its expected value. Also expected utility was used to model comparisons of attitudes of decision makers towards risk. A first decision maker is at least as risk averse as a second if the values $-U''/U'$ for the first decision maker are nowhere smaller than those for the second. Indeed, if then the first decision maker prefers a lottery with nonconstant outcomes (so with risk involved) to a certain amount of money, then the second decision maker will also prefer the risky lottery. This has been found independently by Pratt (1964) and Arrow (1965, 1971).

Still, people have not been fully satisfied by the above results of expected utility. When being informed about the explanation of risk aversion through concavity of the utility function for the first time, most people feel a moment's disappointment. One would think that the utility function rather reflects the decision maker's appreciation of money, than his attitude towards risk. (Be it that the appreciation of money concerns the dealing with risk, which is possibly different from -interval scale = cardinal-appreciations of money in other contexts.) And one would want to reflect attitudes towards risk by means of the way in which probabilities are incorporated, rather than the way in which money is incorporated. An idea is that maybe this can be achieved by transforming probabilities, just as well as Bernoulli transformed payments. Let us denote the transformation by φ . At first sight it may seem self-evident that a lottery is now to be valued by the following modification of (1) :

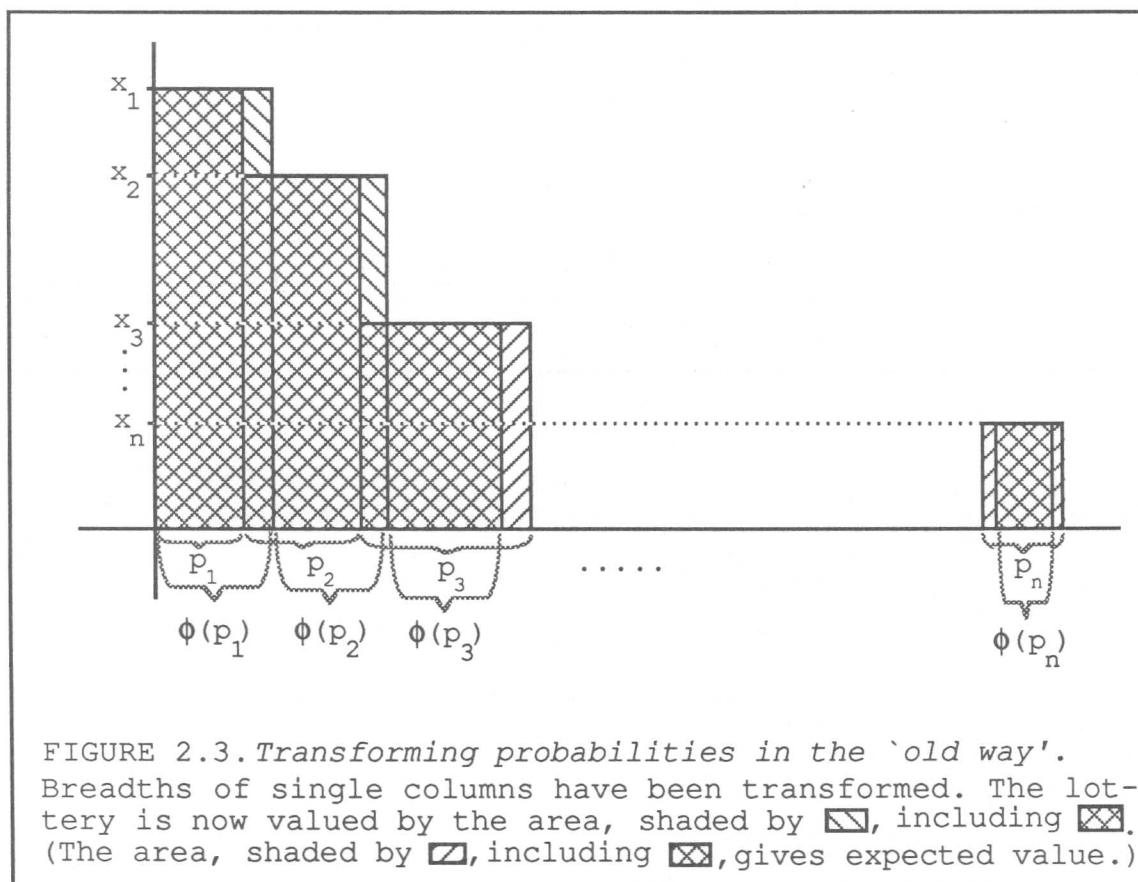
$$\sum_{j=1}^n \varphi(p_j) x_j . \quad (3)$$

It is then natural that φ should be nondecreasing, and that $\varphi(0) = 0$; further we may normalize $\varphi(1) = 1$. This idea was proposed for instance in Edwards (1955, p. 201, second column, lines 34–38, and formulas (3) and (4)); it was considered natural. We shall dispute that naturalness. A first-sight-heuristic argument to suggest that there may be a difference between the payments and the probabilities in (1) is that the payments are free variables, whereas the probabilities are restricted to sum to one. The aim of this paper is to give a more sophisticated heuristic to see the difference.

The new valuation of the lottery of Formula 3 is illustrated in Figure 2.3, by the area shaded by \\\ (including the $\times\times$ -area). Not the heights of the columns, but the breadths, have now been transformed.

The importance of (3) may appear from the central role given to it in 'prospect theory' of Kahneman & Tversky (1979). The idea already occurred in older papers, see for instance Preston & Baratta (1948). Further references are Edwards (1962), Coombs, Bezeminder & Goode (1967), Tversky (1967), and Wallsten (1971). More references are given in Edwards (1954), Luce & Suppes (1965), and Machina (1982, pp. 290–292). Sometimes the transformed probabilities $\varphi(p_j)$ in this approach are called subjective, and the probabilities p_j are called objective. Let us emphasize that the 'subjective probabilities', additive and summing to one, as occurring in Savage (1954), are 'quite different' (Edwards, 1962, p. 109) from the, nonadditive, not-summing-to-one, transformed probabilities. For this reason Edwards (1954) proposed to consider transformed probabilities not as probabilities, but as 'decision weights', a term also used in prospect theory.

We hope to convince the reader that the approach which deals with transformed probabilities as in (3), simply is not very suited.



SECTION 3. THE PROBLEM WITH (3)

In this section we show what the problem is with (3). This has been pointed out for instance in Kahneman & Tversky (1979). Obviously, if the transformation φ is identity, then we deal with decision making as described by older approaches, and no new problems will occur. We shall show that problems occur as soon as φ is not identity. So say φ is not identity. Then, as we shall argue below, there must exist p_1 and p_2 so that :

$$\varphi(p_1 + p_2) \neq \varphi(p_1) + \varphi(p_2) . \quad (4)$$

It is well-known that a function which has equality in (4) for all p_1 and p_2 can only be nonlinear if it is 'very strange'. For instance it can then only be obtained through the 'choice axiom'³ from mathematics, so that it will never be met in practice; its graph must be dense in $[0, 1] \times \mathfrak{R}$, the function is nowhere continuous, is not Lebesgue measurable, and is nowhere monotone (see Aczél, 1966). The latter shows that our nonlinear and monotone φ must indeed for some p_1 and p_2 satisfy the inequality of (4). Suppose

$$\text{Case 1) : } \varphi(p_1 + p_2) > \varphi(p_1) + \varphi(p_2) .$$

To show what may go wrong, we consider what happens if x_1 is decreased in Figure 2.3. That is illustrated in Figure 3.1. At first, not much interesting happens; see Figure 3.1a. The height of the first column has decreased somewhat, and with that

³See Maddy (1988) for its controversial history. It should not be confused with the choice axiom for probabilistic choice theory as in Luce (1959).

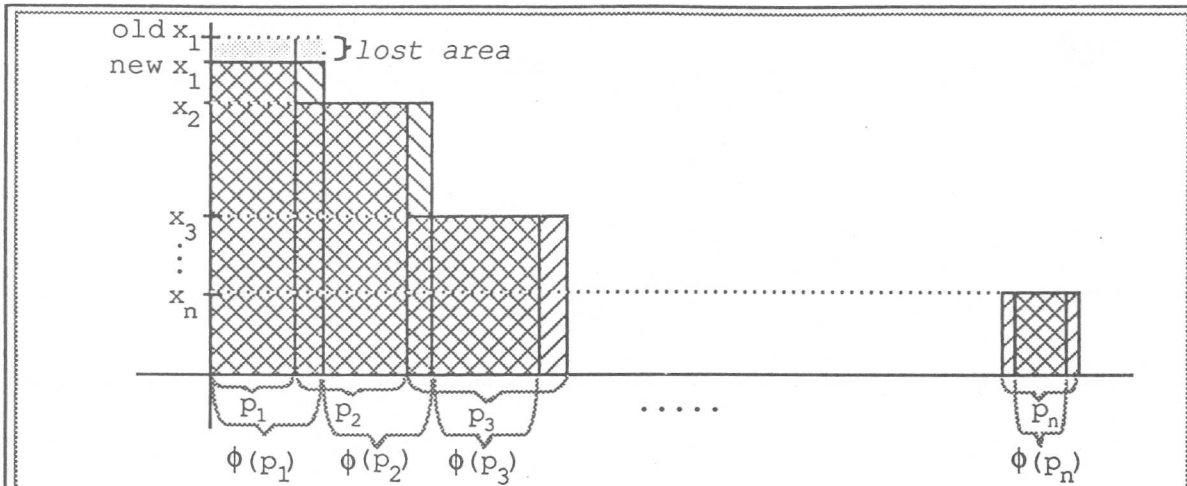


FIGURE 3.1a. Moving x_1 of Figure 2.3 down somewhat. Nothing interesting happens yet. Some area is lost, and the valuation of the lottery decreases somewhat. No breadths of columns are changed.

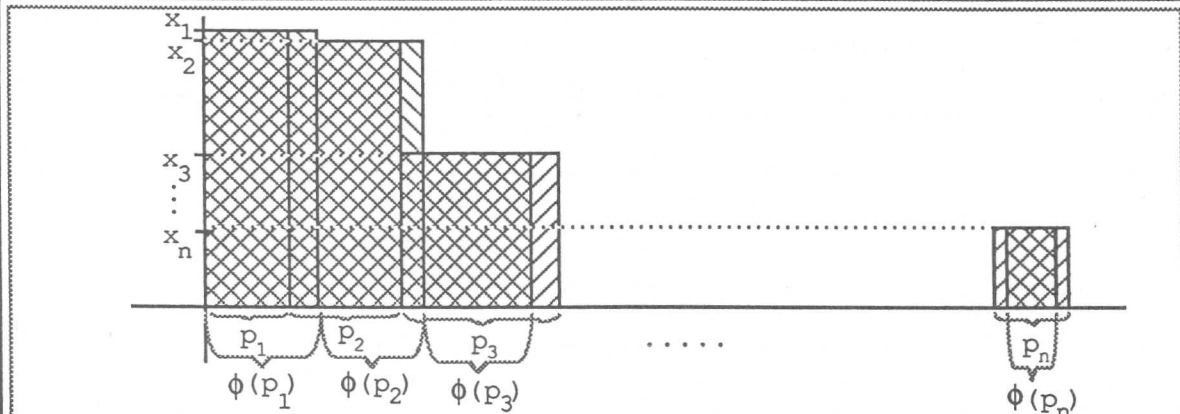


FIGURE 3.1b. Moving x_1 further down. It has gotten very close to x_2 .

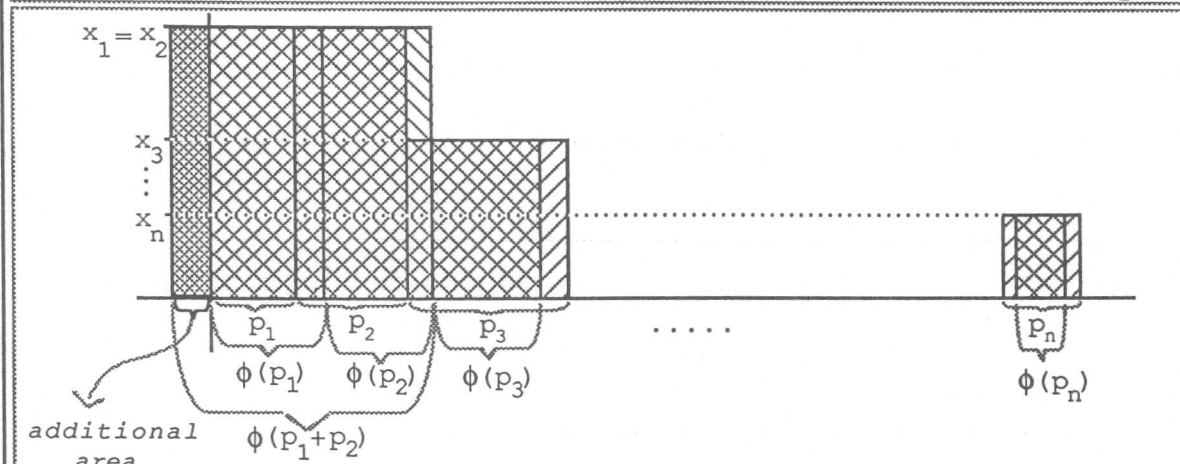


FIGURE 3.1c. x_1 'hits' x_2 . An 'explosion' of new area to the left occurs. It has been shaded denser. We added it to the left for easy comparison. The two columns have been taken together into one broader column. Because of this, the area is larger than in Figure 3.1b. A decrease in payment has led to an increase in area and valuation.

FIGURE 3.1. Formula (3) violates dominance. Applying Formula (3) to different lotteries, differing w.r.t. payment x_1 , reveals violation of stochastic dominance.

its area. Some area has been lost, and the valuation of the lottery has decreased somewhat. Note that the ordering of the outcomes, and the breadths of the columns, have remained unaffected. In Figure 3.1b, x_1 has been decreased further, and with it the area and the valuation of the lottery; x_1 here is very close to x_2 . The problematic phenomenon occurs in Figure 3.1c, where x_1 'hits' x_2 , i.e., $x_1 = x_2$. Then no longer x_1 and x_2 should be treated as different outcomes, with probabilities p_1 respectively p_2 . Now they should be treated as one outcome, with probability $p_1 + p_2$ of occurrence.⁴ The columns with heights x_1 and x_2 should now be taken together into one column, with breadth $\varphi(p_1 + p_2)$. As one sees, this is broader than the two separate columns with joined breadth $\varphi(p_1) + \varphi(p_2)$. So at the moment where x_1 hits x_2 , at the left side an 'explosion' of area occurs. This gives a violation of continuity, which may seem undesirable. However, there is a far more serious problem: it also gives a violation of 'stochastic dominance'! To see this, compare the area, in Figure 3.1b, when x_1 is approaching x_2 so that still $x_1 > x_2$, but x_1 is very close to x_2 , with the area in Figure 3.1c, when $x_1 = x_2$. Because of the 'explosion of area' on the left side, the area in Figure 3.1c is larger. That means that a decrease of payment has induced an increase in valuation! This is exactly what a violation of **(first-order) stochastic dominance** can be defined to be.

In Case 2, where $\varphi(p_1 + p_2) < \varphi(p_1) + \varphi(p_2)$, analogously a violation of stochastic dominance (and continuity) can be demonstrated, now by moving x_2 up to x_1 and by revealing an 'implosion' of area. We do not elaborate this.

The problem that transformed probabilities no longer sum to one has been known for a long time, see for instance Edwards (1954, p. 397/398). Because of this, the breadth of areas like in Figure 3.1 is variable, which can be considered the main cause for the violation of stochastic dominance of formula (3). Nevertheless, it is unknown to us who has been the first to observe explicitly that (3) induces a violation of stochastic dominance.

SECTION 4. ANTICIPATED UTILITY

To introduce anticipated utility, let us first return to the figures of section 2. Note that Figures 2.2 and 2.3 are not fully analogous. In Figure 2.2 the heights all the way down to the x -axis have been transformed, for instance the entire height from x_2 down to the x -axis has been transformed into the height $U(x_2)$. In Figure 2.3 no entire distances to the y -axis have been transformed, but only the breadths of single columns. So naturally the idea of anticipated utility comes to mind: Why not, when transforming breadths/probabilities, also transform entire distances to the y -axis, like it is done in expected utility? This is illustrated in Figure 4.1. First the outcomes are 'rank-ordered' so that $x_1 > x_2 > \dots > x_n$ (in fact only $x_1 \geq x_2 \geq \dots \geq x_n$ is essential). Then the procedure is as follows :

⁴The cruciality of this 'notational' detail in the application of Formula (3) is sometimes felt as unsatisfactory.

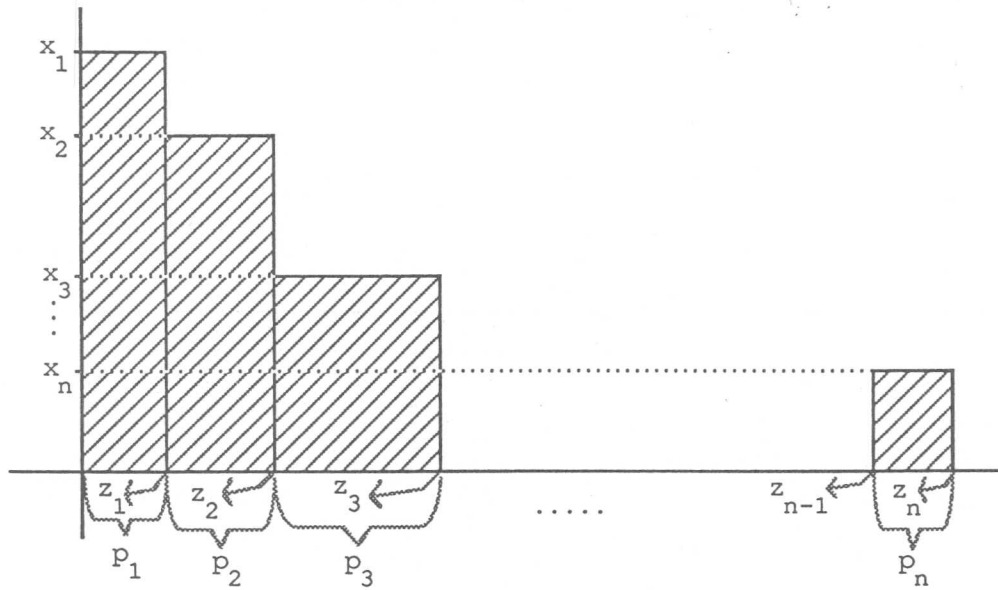


FIGURE 4.1a. *Endpoints of columns.* In the illustration of expected value we have now indicated the points z_1, \dots, z_n , the endpoints of columns. These points will be shifted in Figure 4.1b.

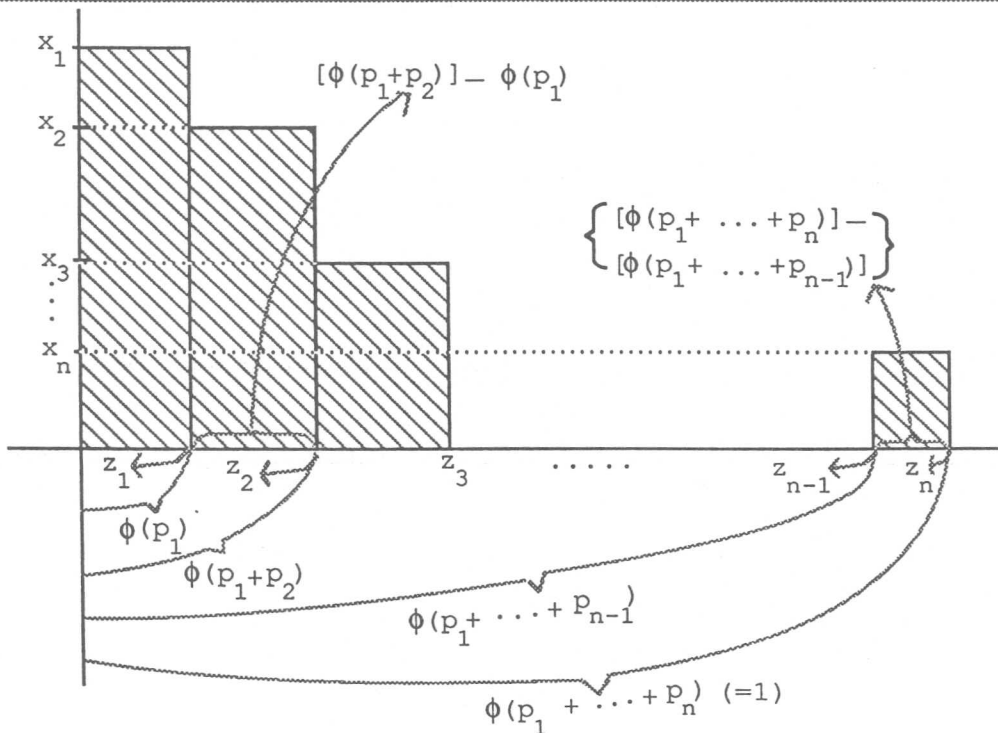


FIGURE 4.1b. *Anticipated utility.* Now entire distances from the points z_j to the y-axis have been transformed. And now z_n remains at a distance 1 from the y-axis.

FIGURE 4.1. *The anticipated utility approach to transformed probabilities.*

Replace p_1 by $\varphi(p_1)$.

So the distance p_1 from z_1 to the y -axis in Figure 4.1a is transformed, in Figure 4.1b, into $\varphi(p_1)$, as it was done in Figure 2.3. The breadth of the first column with height x_1 is now $\varphi(p_1)$.

Replace $p_1 + p_2$ by $\varphi(p_1 + p_2)$.

So the distance $p_1 + p_2$ from z_2 to the y -axis in Figure 4.1a is transformed, in Figure 4.1b, into $\varphi(p_1 + p_2)$, deviating from Figure 2.3. The breadth of the second column with height x_2 is now $\varphi(p_1 + p_2) - \varphi(p_1)$.

·
·
·

Replace $p_1 + p_2 + \dots + p_{n-1}$ by $\varphi(p_1 + p_2 + \dots + p_{n-1})$.

Replace $p_1 + p_2 + \dots + p_n$ by $\varphi(p_1 + p_2 + \dots + p_n)$.

The distance $p_1 + p_2 + \dots + p_n (= 1)$ from z_n to the y -axis in Figure 4.1a is 'transformed', in Figure 4.1b, into $\varphi(p_1 + p_2 + \dots + p_n) (= 1)$, deviating from Figure 2.3. The breadth of the n -th column with height x_n is now $\varphi(p_1 + p_2 + \dots + p_n) - \varphi(p_1 + p_2 + \dots + p_{n-1})$.

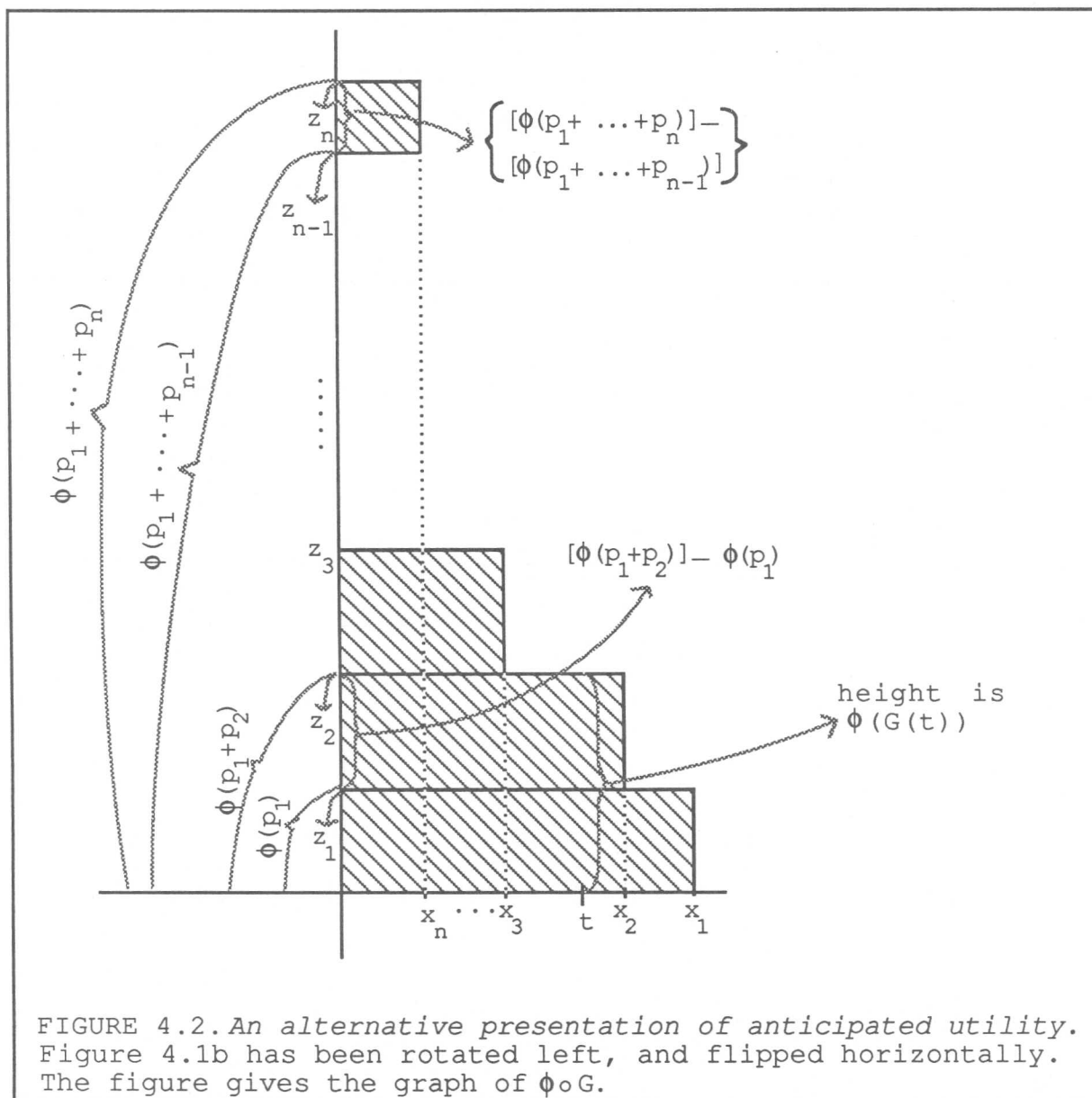
The lottery is now valued by :

$$[\varphi(p_1)]x_1 + [\varphi(p_1 + p_2) - \varphi(p_1)]x_2 + \dots + [\varphi(p_1 + p_2 + \dots + p_n) - \varphi(p_1 + p_2 + \dots + p_{n-1})]x_n . \quad (5)$$

□

This valuation can be reformulated exactly as the valuation by expected value was reformulated in Figure 2.1b. See Figure 4.2. Again, this is obtained from Figure 4.1b by rotating and flipping. Again the size of the shaded area is not affected by these operations. The shape in Figure 4.2 is the graph of the function $\varphi \circ G$, with again G the decumulative distribution function: All heights $G(t)$ have been replaced by their φ -values $\varphi(G(t))$. So the lottery is now valued by

$$\int_{\mathbb{R}_+} \varphi \circ G(t) dt \quad (= [\varphi(p_1)]x_1 + [\varphi(p_1 + p_2) - \varphi(p_1)]x_2 + \dots + [\varphi(p_1 + p_2 + \dots + p_n) - \varphi(p_1 + p_2 + \dots + p_{n-1})]x_n) . \quad (6)$$



SECTION 5. DISCUSSION AND EXAMPLE

The valuation by means of the integral in (6) is exactly the expected value of the lottery which has $\varphi \circ G$ as decumulative distribution function, instead of G .⁵ So one may say that anticipated utility first transforms the decumulative distribution function, and then calculates expected value. Also $\varphi \circ G$ can be considered a ‘capacity’, and the valuation of the lottery the ‘Choquet-integral’ with respect to this capacity. (See for instance Wakker, 1989b, section VI.2 for definitions.) In Wakker (1989d) it is shown how anticipated utility for decision making under risk can be identified with ‘Choquet-expected utility’, the Choquet-integral approach for decision making under

⁵If φ is continuous of the right then indeed $\varphi \circ G$ is a decumulative distribution function of a σ -additive probability distribution.

uncertainty,⁶ for the special case where objective probabilities are known.

To be taken serious, anticipated utility should be able to stand some questions. One of the first questions coming to mind (for instance when one sees the 'main message' in the concluding section 6) is a minor heuristic one: Why should one transform the probabilities of receipt of a fixed outcome or anything *better*, and not the probabilities of receipt of a fixed outcome or anything *worse*? The answer to this question is that it does not matter which of these two is taken. The two methods are 'data-equivalent', i.e., the set of decision strategies that can be modeled by the one method is identical to the set of the other. If one uses a transformation φ in the approach of this paper, then one should use the transformation $\psi : p \mapsto 1 - \varphi(1 - p)$ for the other approach. In (6) one should then replace $\varphi \circ G$ by $\psi \circ F$, with F the usual (cumulative) distribution function. Relations between transformations such as φ and ψ above are central in Gilboa (1989).

Further anticipated utility and the integral in (6) can directly be extended to lotteries with infinite range, and can for instance deal with continuous distributions.

Of more importance is the question of the empirical implications of anticipated utility. We shall not elaborate these. An example of the most important empirical implication will be given in the sequel. Further we confine ourselves to the observation that anticipated utility is free from anomalies in its empirical implications, such as violations of stochastic dominance and of continuity.⁷ Indeed sensible behavioural characterizations of anticipated utility have been obtained, see Quiggin (1982), Schmeidler (1982) and the simultaneous work of Yaari, published in 1987(a); and Gilboa (1987a). Also Wakker (1989a, 1989b Chapter VI) gives a characterization of anticipated utility in its version for decision making under uncertainty; here both transformation of payments and of probabilities are incorporated. For more information on the empirical implications of anticipated utility the reader should consult these references.

The most important empirical implication of expected utility is a kind of 'independence' condition, varying somewhat from context to context. Of course, since anticipated utility is more general than expected utility, its empirical implications will be less restrictive and less strong, so the most important empirical implication of anticipated utility consists of the weakening of (the several versions of) independence to 'comonotonic independence'. Schmeidler (1982) and Wakker (1989c) consider a version of comonotonic independence where probabilities are mixed, Yaari (1987a,b) considers a version of comonotonic independence where payments are mixed. All these versions apply only under the restrictive assumption that there is a kind of linearity w.r.t. the mixtures. We shall not elaborate these conditions, but instead present a version of comonotonic independence that is not restrictive: 'comonotonic coordinate independence'. It was introduced in Wakker (1989a, 1989b Chapter VI), as adaptation of Savage's 'sure-thing principle' (called 'coordinate independence' in Wakker's papers; see end of the following example). In Yaari (1987a,b) and Wakker (1989c, e) intuitive explanations are given in terms of hedging respectively optimism/pessimism. Also Gilboa (1987a) considers a variation of Savage's sure-thing principle, but his variation is more complicated.

⁶In decision making under uncertainty no 'objective' probabilities have to be known in advance; lack of information is modeled through a 'state space' as in Savage (1954).

⁷The latter is a somewhat more technical condition. Still it can add empirical content to other conditions. See Wakker (1988, Example 7.3).

EXAMPLE 1 [*The main empirical implication of the anticipated utility-approach ('comonotonic coordinate independence')*].

Suppose a die will be thrown. Either 1, or 2, . . . , or 6 will come up, each with probability $\frac{1}{6}$. By $x = (x_1, \dots, x_6)$ we denote the 'lottery' giving $\$x_1$ if 1 comes up, . . . , $\$x_6$ if 6 comes up. Suppose a person prefers a bet (x_1, \dots, x_6) to a bet (y_1, \dots, y_6) . Suppose further that $x_1 > x_2 > \dots > x_6$ and that $y_1 > y_2 > \dots > y_6$ ⁸. In anticipated utility the latter is central; to this the term 'comonotonic' refers. Suppose further that x and y have a common outcome, say $x_2 = y_2$. Then, if the person adopts anticipated utility to determine his preferences, **he will not change preference if the common outcome is changed into another common outcome, say some $x'_2 = y'_2$, as long as the change of outcome does not affect the rank-ordering of the outcomes.** I.e., if still $x_1 > x'_2 > \dots > x_6$ and $y_1 > y'_2 > \dots > y_6$, then still $(x_1, x'_2, x_3, x_4, x_5, x_6)$ is preferred to $(y_1, y'_2, y_3, y_4, y_5, y_6)$. Figure 5.1 illustrates how this can be derived.

The term comonotonic, abbreviating 'common monotonic', and introduced by Schmeidler, refers to the fact that the considered lotteries have the same rank-ordering of 'coordinates'. Indeed the involved outcomes can be considered coordinates, associated with fixed results of the throw of the die (such results are called 'states of nature' in decision making under uncertainty); the term 'coordinate' should distinguish this independence condition from the other ('mixture-')versions. 'Coordinate independence' is short for 'independence of equal coordinates'.

If the change of outcome does affect the ordering of outcomes, the preference may be changed. To see this, suppose the person is a ('super-pessimistic') maximin decision maker, i.e., values a lottery by its lowest outcome. This behavior, deviating from expected utility, can be described by anticipated utility, by taking $\varphi(p) = 0$ for all $p < 1$ (then in Figure 4.2 all area to the right of x_n vanishes). Then the above preference holds if $x_6 > y_6$. If $x'_2 (= y'_2)$ becomes smaller than y_6 , so that indeed the rank-ordering of outcomes has been affected, then indeed the preference changes. Then $(x_1, x'_2, x_3, x_4, x_5, x_6)$ becomes equivalent to $(y_1, y'_2, y_3, y_4, y_5, y_6)$, each lottery having value $x'_2 = y'_2$. This phenomenon distinguishes anticipated utility from traditional expected utility. Under expected utility a change of a common outcome can never change the preference (this is the meaning of the 'sure-thing principle'/'coordinate independence'), irrespective of whether or not the rank-ordering of outcomes is affected. \square

Let us take the opportunity to comment on (non-)differentiability of the valuation of lotteries in anticipated utility. The most influential recent deviation of expected utility, available in literature, is Machina (1982). He showed that results concerning ('higher-order') stochastic dominance, and comparisons of risk attitudes, considered before among the most important results of expected utility, could be obtained in a very general framework, not needing expected utility; only a 'Fréchet differentiability' assumption is used. This assumption by itself is technical, but in the presence of other assumptions may get empirical meaning, the status of which is not easy to determine. (Compare Wakker, 1988, Example 7.3.) It is the kind of assumption that is often made without further ado in applied literature. However, as shown by Chew, Karni & Safra (1987, Lemma 1), anticipated utility is *not* Fréchet differentiable. As Karni (1988) pointed out to the author, Figure 4.2 shows this in an easy way. Consider the shift of

⁸ \geq instead of $>$ would suffice

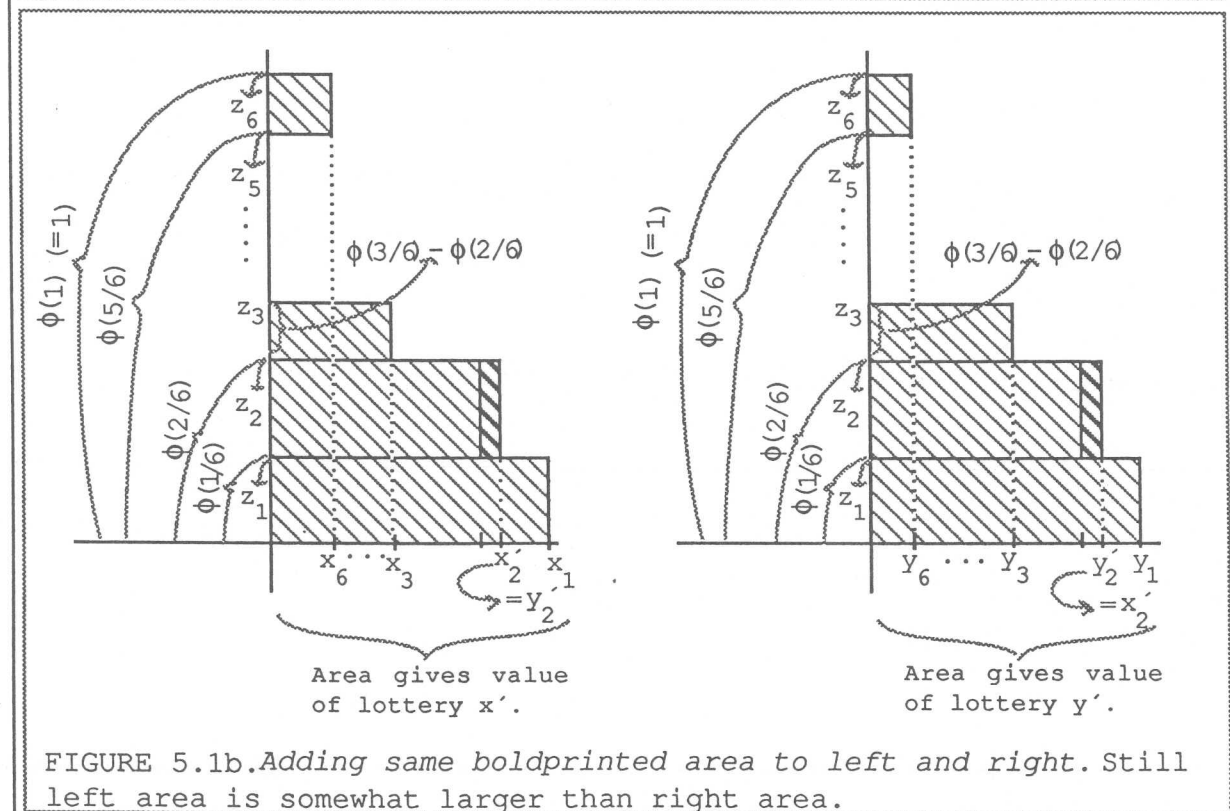
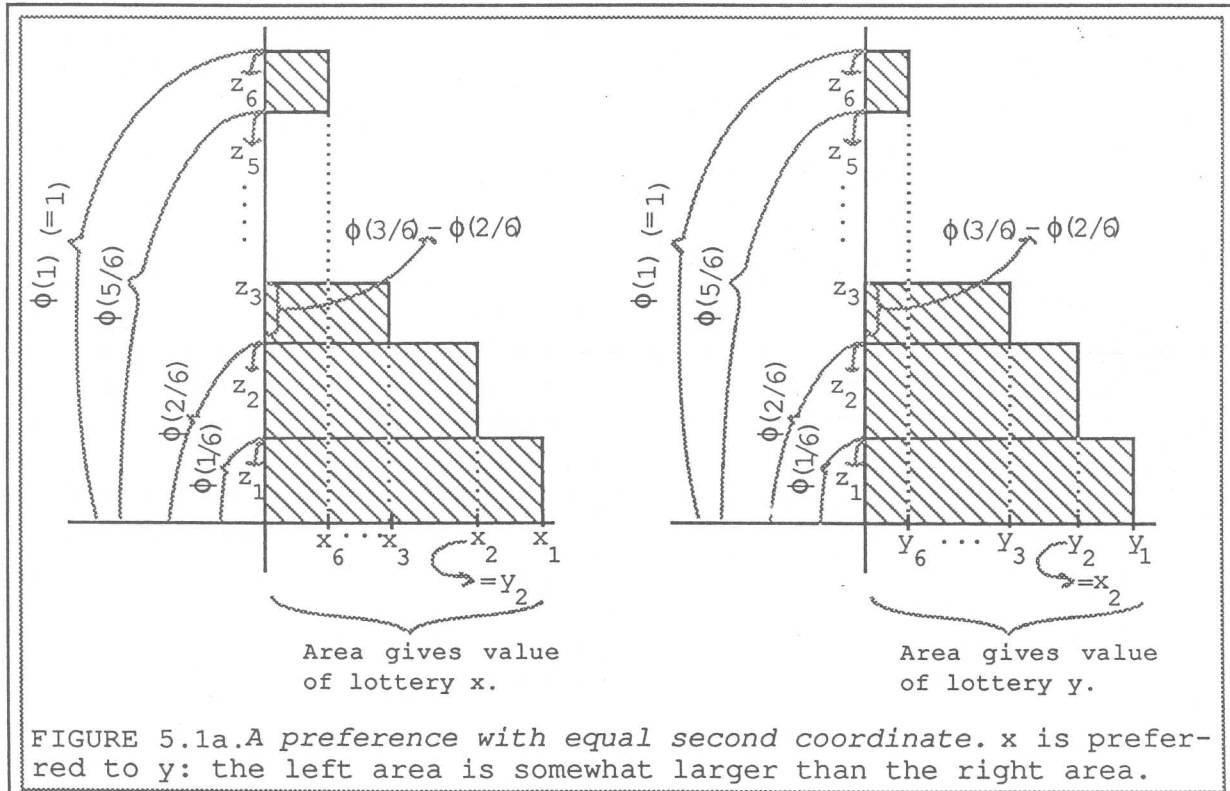


FIGURE 5.1. Comonotonic coordinate independence. The common outcome $x_2=y_2$ from Figure 5.1a is replaced in Figure 5.1b by the common outcome $x'_2=y'_2$ in such a way that the rank-ordering of outcomes has not been affected. By this the areas of the left and right lottery have been increased by the same amount, indicated by \blacksquare . Hence the ordering of the lotteries cannot have been affected, still the left lottery must be preferred to the right.

x_1 towards x_2 . At first this induces a decrease in (Choquet-integral)-value at 'constant speed' $\varphi(p_1)$. However, at the moment that x_1 has 'hit' x_2 and continues becoming smaller than x_2 , the constant speed suddenly changes into the speed $\varphi(p_1 + p_2) - \varphi(p_2)$.⁹ Loosely spoken, for Fréchet differentiability it is necessary that the speed of decrease is the same in all directions, as long as the shifts are infinitesimally small; a sudden change in speed as found above is not allowed. This means that the results of Machina's theory cannot be directly applied to anticipated utility. The above study of Figure 4.2 suggests that 'speed of decrease' is constant for infinitesimal shifts, as long as these shifts go in one direction; this is a loose-hand way to formulate 'Gateaux-differentiability'. Indeed, Chew, Karni & Safra (1987, Corollary 1) show that anticipated utility as presented in this paper¹⁰ is Gateaux-differentiable. We think that the 'sudden change of speed' as observed above, and the implied non-Fréchet differentiability, is one of the most serious intuitive issues of anticipated utility.

SECTION 6. CONCLUSION

The new anticipated utility approach is, in a natural way, dual to expected utility. In the usual expected utility one takes Figure 2.1a, transforms the heights, then takes area. In the new anticipated utility approach one first flips and rotates Figure 2.1a into Figure 2.1b, and only then transforms heights, and takes area. Let us phrase the **main message** of this paper:

If one transforms probabilities, then one should not transform the probabilities of receipt of a fixed outcome, but one should transform the probabilities of receipt of a fixed outcome or anything better.

A final comment: The argumentation of this paper can be no more than heuristic, with arguments based on flippings of figures; this of course depends upon the particular way of illustration chosen in this paper, and the particular arrangement of columns in the Figures. Still we hope that this paper will help the reader getting interested, and encourage him/her to study the firm non-heuristic axiomatic justifications as provided in the references.

ACKNOWLEDGEMENT. Clarifying remarks have been added on the basis of suggestions of Th.G.G. Bezembinder.

REFERENCES

Aczél, J., (1966), '*Lectures on Functional Equations and Their Applications*'. Academic Press, New York.

⁹The latter is most easily seen by comparing it with the 'constant speed of decrease' $\varphi(p_2 + p_1) - \varphi(p_1)$ when in Figure 4.2 in its present version (with $x_1 > x_2$) x_2 is shifted to the left. Also note that the inequality $\varphi(p_1) \neq \varphi(p_1 + p_2) - \varphi(p_2)$ can always be established, according to Formula (4).

¹⁰In fact utility is assumed to be linear, so certainly differentiable, in this paper.

- Allais, M. (1988), 'The General Theory of Random Choices in Relation to the Invariant Cardinal Utility Function and the Specific Probability Function'. In B. R. Munier (Ed.), *Risk, Decision and Rationality*, Reidel, Dordrecht.
- Anger, B. (1977), 'Representations of Capacities', *Mathematische Annalen* 229, 245–258.
- Arrow, K.J. (1965), *Aspects of the Theory of Risk-Bearing*. Academic Bookstore, Helsinki.
- Arrow, K.J. (1971), *Essays in the Theory of Risk-Bearing*. North-Holland, Amsterdam.
- Bernoulli, D. (1738), 'Specimen Theoriae Novae de Mensura Sortis', *Commentarii Academiae Scientiarum Imperialis Petropolitanae* 5, 175–192. Translated into English by L. Sommer (1954), 'Exposition of a New Theory on the Measurement of Risk', *Econometrica* 12, 23–36; or in A.N. Page (Ed., 1968), *Utility Theory: A Book of Readings*, Chapter 11, Wiley, New York.
- Chateauneuf, A. (1988a), 'Uncertainty Aversion and Risk Aversion in Models with Nonadditive Probabilities'. In B. R. Munier (Ed.), *Risk, Decision and Rationality*, 615–629, Reidel, Dordrecht.
- Chateauneuf, A. (1988b), 'Decomposable Measures, Distorted Probabilities and Concave Capacities', Working paper, Groupe de Mathématiques Economiques, Université de Paris I.
- Chateauneuf, A. & J.-Y. Jaffray (1987), 'Some Characterizations of Lower Probabilities and Other Monotone Capacities', *Mathematical Social Sciences*, forthcoming.
- Chew, S.H. (1985), 'An Axiomatization of the Rank-Dependent Quasilinear Mean Generalizing the Gini Mean and the Quasilinear Mean', Economics Working Paper # 156, Johns Hopkins University.
- Chew, S.H., E. Karni, & Z. Safra (1987), 'Risk Aversion in the Theory of Expected Utility with Rank Dependent Probabilities', *Journal of Economic Theory* 42, 370–381.
- Choquet, G. (1953-4), 'Theory of Capacities', *Annales de l'Institut Fourier* (Grenoble) 5, 131–295.
- Coombs, C.A., Th.G.G. Bezembinder, & F.M. Goode (1967), 'Testing Expectation Theories without Measuring Utility or Subjective Probability', *Journal of Mathematical Psychology* 4, 72–103.
- Dellacherie, C. (1970), 'Quelques Commentaires sur les Prolongements de Capacités', *Seminaire de Probabilités V Strasbourg*, (Lecture Notes in Mathematics 191), Springer Verlag, Berlin.
- Denneberg, D. (1988), 'On Non-Expected-Utility Preferences', paper presented at 4th FUR conference, Budapest, 1988.
- Dubois, D. & H. Prade (1988), 'Modelling Uncertainty and Inductive Inference: A Survey of Recent Non-Additive Probability Systems', *Acta Psychologica* 68, forthcoming.
- Ebert, U. (1987), 'Measurement of Inequality: An Attempt at Unification and Generalization', *Social Choice and Welfare*, forthcoming.
- Ebert, U. (1988), 'Rawls and Bentham Reconciled', *Theory and Decision* 24, 215–223.
- Edwards, W. (1954), 'The Theory of Decision Making', *Psychological Bulletin* 51, 380–417.
- Edwards, W. (1962), 'Subjective Probabilities Inferred from Decisions', *Psychological Review* 69, 109–135.

- Fellner, W. (1961), 'Distortion of Subjective Probabilities as a Reaction to Uncertainty', *Quarterly Journal of Economics* 75, 670–690.
- Fishburn, P.C. (1988), *Nonlinear Preference and Utility Theory*. Johns Hopkins University Press, Baltimore.
- Gilboa, I. (1985), 'Subjective Distortions of Probabilities and Non-Additive Probabilities', Working paper 18–85, Foerder Institute for Economic Research, Tel-Aviv University, Ramat Aviv, Israel.
- Gilboa, I. (1987a), 'Expected Utility with Purely Subjective Non-Additive Probabilities', *Journal of Mathematical Economics* 16, 65–88.
- Gilboa, I. (1987b), 'Expectations and Variations in Multi-Period Decisions', *Econometrica*, forthcoming. Working paper 10–87, Foerder Institute for Economic Research, Tel-Aviv University, Ramat Aviv, Israel.
- Gilboa, I. (1989), 'Duality in Non-Additive Expected Utility Theory'. In P.C. Fishburn & I.H. LaValle (Eds.), *Choice under Uncertainty*, Annals of Operations Research, J.C. Baltzer AG., Basel, forthcoming.
- Hilton, R.W. (1988), 'Risk Attitude under Two Alternative Theories of Choice under Risk', *Journal of Economic Behaviour and Organization* 9, 119–136.
- Huber, P.J. (1981), *Robust Statistics*. Wiley, New York.
- Kahneman, D. & A. Tversky (1979), 'Prospect Theory: An Analysis of Decision under Risk', *Econometrica* 47, 263–291.
- Karni, E. (1988), Personal communication.
- Krantz, D.H., R.D. Luce, P. Suppes, & A. Tversky (1971), *Foundations of Measurement, Vol. I. (Additive and Polynomial Representations)*. Academic Press, New York.
- Luce, R.D. & P. Suppes (1965), 'Preference, Utility, and Subjective Probability'. In R.D. Luce, R.R. Bush, & E. Galanter (Eds.), *Handbook of Mathematical Psychology*, III, Wiley, New York.
- Machina, M.J. (1982), 'Expected Utility' Analysis without the Independence Axiom', *Econometrica* 50, 277–323.
- Machina, M.J. (1987), 'Choice under Uncertainty: Problems Solved and Unsolved', *Economic Perspectives* 1, 121–154.
- Maddy, P. (1988), 'Believing the Axioms. I', *The Journal of Symbolic Logic* 53, 481–511.
- Nakamura, Y. (1989), 'Subjective Expected Utility with Non-Additive Probabilities on Finite State Space', Discussion paper No. 400, April 1989, University of Tsukuba, Tsukuba, Ibaraki, Japan.
- Pratt, J.W. (1964), 'Risk Aversion in the Small and in the Large', *Econometrica* 32, 122–136.
- Preston, M.G. & P. Baratta (1948), 'An Experimental Study of the Auction Value of an Uncertain Outcome', *American Journal of Psychology* 61, 183–193.
- Quiggin, J. (1982), 'A Theory of Anticipated Utility', *Journal of Economic Behaviour and Organization* 3, 323–343.
- Röell, A. (1987), 'Risk Aversion in Quiggin and Yaari's Rank-Order Model of Choice under Uncertainty', (Supplement to the) *Economic Journal* 97, 143–160.
- Savage, L.J. (1954), *The Foundations of Statistics*. Wiley, New York. (Second edition 1972, Dover, New York.)
- Schlee, E. (1988), 'The Value of Information in Anticipated Utility Theory', University of Alabama, Department of Economics, October 1988.

- Schmeidler, D. (1982), 'Subjective Probability without Additivity', Foerder Institut of Economic Research, Tel Aviv University, Tel Aviv, Israel. (Rewritten as Schmeidler, 1984.)
- Schmeidler, D. (1984), 'Nonadditive Probabilities and Convex Games'. Caress working paper 84-21 (second part), University of Pennsylvania, Center for Analytic Research in Economics and the Social Sciences, Philadelphia, PA.
- Schmeidler, D. (1986), 'Integral Representation without Additivity'. *Proceedings of the American Mathematical Society* 97, 255-261.
- Segal, U. (1988), 'Anticipated Utility: A Measure Representation Approach', Working paper 8803, University of Toronto, Department of Economics and Institute for Policy Analysis, Toronto, Canada. Submitted to *Annals of Operations Research*. Rewritten version of Segal (1984), 'Nonlinear Decision Weights with the Independence Axiom', Working paper 353, University of California, Department of Economics, Los Angeles, USA.
- Tversky, A. (1967), 'Additivity, Utility, and Subjective Probability', *Journal of Mathematical Psychology* 4, 175-201.
- von Neumann, J. & O. Morgenstern (1944, 1947, 1953), '*Theory of Games and Economic Behavior*'. Princeton University Press, Princeton NJ.
- Wakker, P.P. (1987), 'From Decision Making under Uncertainty to Game Theory'. In H.J.M. Peters & O.J. Vriete (Eds.), *Surveys of Game Theory and Related Topics*, 163-180, CWI Tract 39, Centre for Mathematics and Computer Science, Amsterdam.
- Wakker, P.P. (1988), 'The Algebraic Versus the Topological Approach to Additive Representations', *Journal of Mathematical Psychology* 32, 421-435.
- Wakker, P.P. (1989a), 'Continuous Subjective Expected Utility with Nonadditive Probabilities', *Journal of Mathematical Economics* 18, 1-27.
- Wakker, P.P. (1989b), '*Additive Representations of Preferences, A New Foundation of Decision Analysis*'. Kluwer Academic Publishers, Dordrecht.
- Wakker, P.P. (1989c), 'A Behavioral Foundation for Fuzzy Measures', Internal report 89 MA 01, University of Nijmegen, NICI, Department of Mathematical Psychology, Nijmegen, The Netherlands. Submitted to *Fuzzy Sets and Systems*.
- Wakker, P.P. (1989d), 'Stochastic Dominance Implies the Equality [Choquet-Expected Utility = Anticipated Utility]', Internal report 89 MA 02, University of Nijmegen, NICI, Department of Mathematical Psychology, Nijmegen, The Netherlands.
- Wakker, P.P. (1989e), 'Characterizing Optimism and Pessimism Directly through Comonotonicity', Internal report 89 NICI 06, University of Nijmegen, NICI, Department of Mathematical Psychology, Nijmegen, The Netherlands.
- Wallsten, T. (1971), 'Subjective Expected Utility Theory and Subjects' Probability Estimates: Use of Measurement-Free Techniques', *Journal of Experimental Psychology* 88, 31-40.
- Wallsten, T.S. & B.H. Forsyth (1985), 'On the Usefulness, Representation, and Validation of Non-Additive Probability Judgements for Risk Assessment', Department of Psychology, University of North Carolina, Chapel Hill, NC, USA.
- Yaari, M.E. (1987a), 'The Dual Theory of Choice under Risk', *Econometrica* 55, 95-115.
- Yaari, M.E. (1987b), 'Univariate and Multivariate Comparisons of Risk Aversion: a New Approach'. In W.P. Heller, R.M. Starr, & D.A. Starrett (Eds.), *Uncertainty, Information and Communication, Essays in honor of Kenneth J. Arrow*, Vol. III,

173–187, Cambridge University Press, Cambridge.

Yaari, M.E. (1988), 'A Controversial Proposal Concerning Inequality Measurement', *Journal of Economic Theory* 44, 381–397.

Yager, R.R. (1988), 'On Ordered Weighted Averaging Aggregation Operators in Multicriteria Decisionmaking', *IEEE Transactions on Systems, Man, and Cybernetics* 18, 183–190.