On the Intuition of Rank-Dependent Utility

ENRICO DIECIDUE AND PETER P. WAKKER*
wakker@fee.uva.nl
CentER, Tilburg University, P.O. Box 90153, Tilburg, 5000 LE, The Netherlands

Abstract

Among the most popular models for decision under risk and uncertainty are the rank-dependent models, introduced by Quiggin and Schmeidler. Central concepts in these models are rank-dependence and comonotonicity. It has been suggested that these concepts are technical tools that have no intuitive or empirical content. This paper describes such contents. As a result, rank-dependence and comonotonicity become natural concepts upon which preference conditions, empirical tests, and improvements in utility measurement can be based. Further, a new derivation of the rank-dependent models is obtained. It is not based on observable preference axioms or on empirical data, but naturally follows from the intuitive perspective assumed. We think that the popularity of the rank-dependent theories is mainly due to the natural concepts used in these theories.

Keywords: rank-dependence, comonotonicity, Choquet integral, pessimism, uncertainty aversion, prospect theory

JEL Classification: D8i, C60

Many models for decision under risk and uncertainty have been proposed that deviate from classical expected utility. Among the most popular are the rank-dependent models. They were introduced by Quiggin (1981) for decision under risk (known probabilities) and by Schmeidler (1989) for decision under uncertainty (unknown probabilities). Rank-dependence has been incorporated in original prospect theory (Kahneman and Tversky, 1979), leading to cumulative prospect theory (Tversky and Kahneman, 1992). The present paper proposes an intuitive justification of rank-dependence, building on Lopes (1984), Weber (1994), and Yaari (1987). A new derivation of rank-dependent utility is presented that naturally follows from the intuitive conditions. For intuitive arguments for betweenness models, see Epstein (1992).

In order to generate fruitful applications, a decision model should satisfy three requirements. First, it should be mathematically sound. For instance, it should not exhibit behavioral anomalies such as implausible violations of stochastic dominance (Fishburn, 1978). This first requirement can be guaranteed by preference axiomatizations. For the rank-dependent models, such axiomatizations were given by Quiggin (1982), Schmeidler...

The second requirement for a decision model concerns its empirical performance. It has been found that rank-dependent utility can accommodate several empirical violations of expected utility. The study of its empirical potential is still going on today (Abdellaoui and Munier, 1999; Birnbaum and Mcintosh, 1996; Bleichrodt and Pinto, 2000; Gonzalez and Wu, 1999; Harless and Camerer, 1994; Tversky and Fox, 1995).

The third requirement is that the model should be intuitively plausible. Its concepts should provide new insights and be economically meaningful. Future connections with concepts from other fields should be conceivable (nomological validity). Only a few authors have given intuitive arguments for rank-dependence. These arguments are scattered around over various papers in different fields. It has been suggested recently that a complete intuitive foundation of rank-dependence is still lacking (Luce, 1996a, p. 85; Luce, 1996b, p. 304; Safra and Segal, 1998, p. 28). Providing such a foundation is the purpose of this paper. We will argue, using the terminology of Backhouse (1998, p. 1857), that rank-dependence relates to real-world (psychological) concepts. As suggested by Backhouse, such arguments are, “in the last resort, informal” (see also Loomes and Sugden, 1982, p. 817).

The paper is structured as follows. Section 1 presents the first attempt to model non-additive probabilities, commonly used before the 1980s. Section 2 describes the intuition of rank-dependence for decision under risk. The rank-dependent utility formula follows from this intuition in a natural and elementary manner (Section 3). The intuition also leads to natural ways of modeling pessimism and optimism, two common attitudes towards probabilistic risk (Section 4). Section 5 extends the foundation to uncertainty. It shows that Quiggin’s (1981) contribution for risk and Schmeidler’s (1989) contribution for uncertainty can be based on the same intuition. Using the intuitive foundation of the preceding sections, Section 6 argues that preference conditions and measurement procedures based on the comonotonicity restriction are not only valid under the rank-dependent theory but also have merits in the real world. Conclusions and comments are given in Section 7. Appendix A discusses some intuitive arguments for rank-dependence that were presented before in the literature and Appendix B gives proofs.

1. The first attempt

Consider a lottery \( (p_1, x_1; \ldots; p_n, x_n) \), yielding outcome \( x_j \) with probability \( p_j \), \( j = 1, \ldots, n \). The probabilities \( p_1, \ldots, p_n \) are nonnegative and sum to one. In this paper, outcomes are real numbers designating money. It is assumed throughout that the lottery is evaluated by the following formula, called the general weighting model:

\[
\sum_{j=1}^{n} \pi_j U(x_j).
\]  

(1)

\( U \) is the utility function and the \( \pi_j \)'s are called decision weights. The decision weights are nonnegative and sum to one, and will be discussed later. The general weighting model
is not intended to immediately imply operational predictions but serves as a general point of departure. Intuitive arguments will be formulated in terms of the model, and operational implications will be established subsequently.

Stochastic dominance is assumed throughout the paper. It means that moving positive probability mass from an outcome to a strictly higher outcome leads to a strictly higher evaluation. This assumption implies that the utility function is strictly increasing (as follows from considering riskless lotteries). We do not yet make any further assumption about the decision weights, and they may depend on the entire lottery for now. In a descriptive context, \( \pi_j \) can be interpreted as the attention given to outcome \( x_j \), possibly due to misperception of probability. In a normative context, \( \pi_j \) can be interpreted as an importance weight for outcome \( x_j \) that may deliberately have been chosen different than the probability \( p_j \).

It may be possible to relate decision weights to psychological notions such as the time span during which the decision maker looks at outcomes (Johnson and Schkade, 1989). An empirical operationalization of decision weights is, however, not our purpose at this stage. When further assumptions have been added, the decision weights will become operational.

Utility is assumed to be independent of the lottery under consideration. Like decision weights, utility is not operational at this stage but will become so later when further assumptions have been added. Utility can be operationalized if it is interpreted in the riskless sense of Allais (1953). Expected utility is the special case of the general weighting model where \( \pi_j \) agrees with \( p_j \) for all \( j \).

As a preparation for what follows, and for historical reasons, we start with the following assumption. It will turn out to be too restrictive for our purposes and will be relaxed later on.

**Assumption 1** [independence of beliefs from tastes]. The decision weight \( \pi_j \) of receiving outcome \( x_j \) depends only on the probability \( p_j \).

The assumption requires that the decision weight \( \pi_j \) is independent of the outcomes and the other probabilities of the lottery. We can now write \( w(p) \) for the decision weight generated by a probability \( p \), thus defining a function \( w \). The general weighting model becomes

\[
\sum_{j=1}^{n} w(p_j)U(x_j).
\]

As we will see next, the requirement that the \( w(p_j) \)s sum to one implies expected utility. The proof is given in Appendix B.

**Theorem 2.** Under Assumption 1, the general weighting model (1) reduces to expected utility, i.e. \( w(p) = p \).

To obtain Eq. 2 with a nonlinear \( w \) function, the requirement that decision weights sum to one will have to be relaxed. This was indeed the approach originally taken in the literature (Edwards, 1955; Preston and Baratta, 1948). The resulting model, however,
leads to violations of stochastic dominance (Fishburn, 1978). We conclude that a transformation of probabilities, independently of outcomes, is not well possible. To obtain a decision theory with transformed probabilities, an additional relaxation of the expected utility principles is required. Such a relaxation, rank-dependence, was introduced by Quiggin (1981). Its intuition is explained in the next section and the rest of the paper elaborates on this intuition.

2. The intuition and definition of rank-dependence for decision under risk

The intuition of rank-dependence entails that the attention given to an outcome depends not only on the probability of the outcome but also on the favorability of the outcome in comparison to the other possible outcomes. To illustrate this intuition, assume that the decision maker is a pessimist and evaluates the lottery \( \left( \frac{1}{3}, 30; \frac{1}{3}, 20; \frac{1}{3}, 10 \right) \). Then he will pay more than \( \frac{1}{3} \) of his attention to outcome 10, the reason being that 10 is the worst outcome. Say that \( \pi_1 \), the decision weight for outcome 10, is \( \frac{1}{3} \). The decision maker, accordingly, pays relatively less attention to each of the other outcomes (\( \pi_2 = \frac{1}{3} \) if \( \pi_3 = \frac{1}{3} \)). Being a pessimist, he will pay more than half of the remaining attention to outcome 20 and, hence, \( \pi_2 > \frac{1}{3} \); say \( \pi_2 = \frac{1}{2} \). The remainder of the attention, devoted to outcome 30, is small \( \pi_1 = \frac{1}{6} \). Next consider the lottery with outcome 20 changed into 0, i.e. \( \left( \frac{1}{3}, 30; \frac{1}{3}, 0; \frac{1}{3}, 10 \right) \). The outcome 10 is no longer the worst outcome and a pessimist will therefore pay less attention to it than in the first lottery. In human behavior, such attitudes are commonly observed in every-day life. Rank-dependence is a psychologically realistic phenomenon. Savage (1954, end of Chapter 4) already pointed out that there is no room for expressing optimism or pessimism in traditional expected utility.

Descriptively, a pessimistic attitude can result from an irrational belief that unfavorable events tend to happen more often, leading to an unrealistic overweighting of unfavorable likelihoods (Murphy's law). If rank-dependence is taken normatively, then a pessimistic attitude can result from conscious and deliberate decisions. The decision maker may decide that unfavorable outcomes are especially important in decision making and therefore should receive more attention than equally likely favorable outcomes (Ellsberg, 1961, p. 667; Fellner, 1961, p. 681; Lopes and Oden, 1999, p. 310; Weber, 1994, p. 236).

Empirically, another kind of rank-dependence is often found, where subjects pay much attention not only to the worst outcomes but also to the best outcomes. Less attention is paid to the intermediate outcomes. This phenomenon may result from extreme outcomes being more noticeable. It once more illustrates the realistic nature of rank-dependence. A discussion of the phenomenon is given in Section 4.

Further generalizations of expected utility could obviously be considered. To some extent, the decision weight of an outcome will depend not only on whether it is better than some other outcome but also on how much better it is. Such generalizations may be considered in future developments. It should, however, be kept in mind that a theory should not be too general. The theory should be sufficiently restrictive to allow for specific predictions. In this sense, rank-dependence can be considered a pragmatic compromise between generality and parsimony. It incorporates some major deviations
ON THE INTUITION OF RANK-DEPENDENT UTILITY

1. Ranking

Assumption 3 [rank-dependence]. The decision weight \( \pi_j \) of receiving outcome \( x_j \) depends only on its probability \( p_j \) and its ranking position.

This assumption has relaxed Assumption 1 by also permitting rank-dependence. To illustrate the assumption, consider the lottery \((\frac{1}{3}, 30; \frac{1}{3}, 20; \frac{1}{3}, 10)\). The ranking position of outcome 10 is \( \frac{1}{3} \). For the lottery \((\frac{1}{3}, 25; \frac{1}{3}, 12)\), the ranking position of outcome 12 is also \( \frac{1}{3} \). The two outcomes also have the same probability. By Assumption 3, they must have the same decision weight.

3. Operational implications: rank-dependent utility for risk

With Assumption 3 added, the decision weights become operational and empirical predictions can be derived from the decision weights. For example, with ~ denoting equivalence, assume that

\[(p_1, 10; p_2, 2; p_3, 1) \sim (q_1, 12; q_2, 2; q_3, 0).\]
Then the decision weight of outcome 2 in the left lottery exceeds the corresponding decision weight in the right lottery if and only if, with \( \succ \) denoting preference,

\[
(p_1, 10; p_2, 3; p_3, 1) \succ (q_1, 12; q_2, 3; q_3, 0).
\]

The claim follows because, under Assumption 3, the middle outcomes of the left lotteries (2 in the upper lottery and 3 in the lower) have the same decision weight \( \pi_2 \), and the middle outcomes of the right lotteries (2 in the upper lottery and 3 in the lower) have the same decision weight \( \pi'_2 \). The increase in evaluation of the left lottery, \( \pi_2(U(3) - U(2)) \), apparently exceeds the increase in evaluation of the right lottery, \( \pi'_2(U(3) - U(2)) \). Consequently, \( \pi_2 \geq \pi'_2 \). That is, the decision weights show where to put your money (see Sarin and Wakker, 1998, using an idea of Gilboa, 1987).

We next demonstrate that rank-dependent utility follows from the general weighting model and Assumption 3. Assumption 3 implies in particular that the decision weight of a maximal outcome of a lottery depends only on its probability \( p \), its ranking position always being one. The function \( w(p) \) is defined as this decision weight. Let us emphasize that \( w(p) \) is the decision weight generated by the probability \( p \) when associated with the best outcome. Obviously, \( w(0) = 0, w(1) = 1 \), and \( w \) is strictly increasing because of stochastic dominance.

The general rank-dependent formula for the lottery \( (p_1, x_1; \ldots; p_n, x_n) \) with \( x_1 > \cdots > x_n \) can be expressed in terms of the function \( w \). The decision weight \( \pi_i \) is equal to \( w(p_i) \) by definition. We next turn to the decision weight of outcome \( x_i \) for some general \( i \). The following observation serves as a preparation.

**Observation.** The total decision weight assigned to outcomes \( x_1, \ldots, x_i \), i.e. \( \pi_1 + \cdots + \pi_i \), is \( w(p_1 + \cdots + p_i) \).

**Explanation.** Consider the lotteries \( (p_1, x_1; \ldots; p_i, x_i; p_{i+1}, x_{i+1}; \cdots; p_n, x_n) \) and \( ((p_1 + \cdots + p_i), z; p_{i+1}, x_{i+1}; \cdots; p_n, x_n) \) for any outcome \( z \) exceeding \( x_{i+1} \), e.g., \( z = x_1 \). Because decision weights must sum to one, \( \pi_1 + \cdots + \pi_i = 1 - \pi_{i+1} - \cdots - \pi_n = w(p_1 + \cdots + p_i) \), where the second equality can be inferred from inspecting the second lottery. Crucial in this explanation is that the outcomes \( x_{i+1}, \ldots, x_n \) all have the same ranking position in the two lotteries and therefore, by Assumption 3, the same decision weights denoted by \( \pi_{i+1}, \ldots, \pi_n \).

The decision weight \( \pi_i \) of outcome \( x_i \) is \( \pi_1 + \cdots + \pi_i - (\pi_1 + \cdots + \pi_{i-1}) \). By the preceding observation, \( \pi_i \) is equal to \( w(p_1 + \cdots + p_i) - w(p_1 + \cdots + p_{i-1}) \). Therefore, every decision weight can be expressed in terms of \( w \). In agreement with the rank-dependence Assumption 3, the decision weight of \( x_i \) depends only on its probability \( p_i \) and its ranking position \( q = p_1 + \cdots + p_{i-1} \), because it can be written as \( w(p_i + 1 - q) - w(1 - q) \).

Let us summarize. The model that has been derived is called rank-dependent utility (RDU). If \( x_1 > \cdots > x_n \) then

\[
RDU(p_1, x_1; \cdots; p_n, x_n) = \sum_{j=1}^{n} \pi_j U(x_j)
\]
where, for each \( j \),

\[
\pi_j = w(p_1 + \cdots + p_j) - w(p_1 + \cdots + p_{j-1}).
\]

In particular, \( \pi_1 = w(p_1) \).

**Conclusion 4.** Under the general weighting model (Eq. 1), stochastic dominance and Assumption 3 imply rank-dependent utility.

The preceding analysis used the function \( w(p) \), the decision weight generated by probability \( p \) when associated with the best outcome. An equivalent analysis could have been presented in terms of a dual function \( w^*(p) \), describing the decision weight generated by probability \( p \) when associated with the worst outcome. The two functions are dual in the sense that \( w^*(p) = 1 - w(1 - p) \) for all \( p \). This duality follows because the decision weights should sum to one for any lottery \((p, M; 1 - p, m)\) with outcomes \( M > m \). The analysis can be based both on \( w \) and on \( w^* \), but it should be kept in mind whether the function describes decision weights of best outcomes or of worst outcomes. In (3), \( \pi_j \) can as well be expressed in terms of \( w^* \), \( \pi_j = w^*(p_j + \cdots + p_n) - w^*(p_{j+1} + \cdots + p_n) \) for each \( j \). \( w \) can be called the goodnews weighting function and \( w^* \) the badnews weighting function.

The decision weights are now uniquely determined and can be derived from observable choice. Most empirical studies of decision weights have used simultaneous parametric fittings for \( U \) and \( w \). Non-parametric fittings still involving utility estimation were provided by three independent and simultaneous studies: Abdellaoui (2000), Bleichrodt and Pinto (2000), and Gonzalez and Wu (1999). Abdellaoui (1999) introduced a parameter-free method for measuring decision weights without the need to estimate utilities.

Other nonexpected utility models than the rank-dependent ones can be derived from the general weighting model. For example, if the decision weights do not depend on the rank-ordering of outcomes but instead on the equivalence class that a lottery is contained in, then betweenness models result (Chew, 1989; Epstein, 1992). These models are outside the scope of this paper.

We hope that the preceding explanation has demonstrated that RDU is not solely a mathematical device for deriving decisions from nonlinear probabilities. The theory is based on two intuitive assumptions regarding decision making. First, people process probabilities in a nonlinear manner. Second, the attention people pay to outcomes depends on how good or bad these outcomes are. The RDU formula naturally follows from these two intuitive assumptions.

4. **Pessimism and optimism**

This section shows how rank-dependence can describe phenomena outside the domain of expected utility. We first consider pessimism. Assume that a lottery yields outcome \( x \) with probability \( p \). Let \( q \) denote the ranking position of \( x \), i.e. the probability of receiving
a lower or equal outcome. The decision weight of \( x \) then is \( w(p + (1 - q)) - w(1 - q) \). Under pessimism, improving the ranking position (increasing the probability \( q \) of receiving something not preferred) decreases the decision weight of \( x \). It is well-known that \( w(p + (1 - q)) - w(1 - q) \) is decreasing in \( q \) if and only if \( w \) is convex. Hence, pessimism is characterized by a convex weighting function.

Similarly, optimism corresponds to a decision weight \( w(p + (1 - q)) - w(1 - q) \) that is increasing in \( q \), and thus to a concave weighting function. This rank-dependent way of modeling pessimism and optimism was suggested before by Quiggin (1982, p. 335). It was described in full by Yaari (1987, p. 108) and, subsequently, by many other authors. It is in full agreement with the intuition advanced in this paper. Similar effects have been demonstrated in other contexts (Viscusi, 1997, p. 1667).

In empirical investigations, many observed weighting functions are not completely convex or concave but exhibit a mixed pattern. They are concave for small probabilities and convex for moderate and high probabilities. This pattern is called inverse-S. It implies that subjects pay much attention to best and worst outcomes, and little attention to intermediate outcomes (Quiggin, 1982; Weber, 1994). Its empirical support is reviewed by Wakker (2001). Counterevidence can be found in Birnbaum and McIntosh (1996) and Birnbaum and Navarrete (1998). For a psychological theory about the attention paid to low outcomes (security) and high outcomes (potential), see Lopes and Oden (1999). The pattern suggests that people are overly sensitive to changes from impossible to possible and from possible to certain but are insufficiently sensitive to probabilistic information otherwise (Karmarkar, 1978; Tversky and Wakker, 1995).

The inverse-S shape predicts that people are optimistic and, hence, risk seeking for gambles that yield gains with small probabilities such as found in public lotteries. People are pessimistic and, hence, risk averse for gambles that yield losses with small probabilities, which is relevant for insurance. The simultaneous existence of gambling and insurance, a classical paradox in economics, can therefore be explained by the inverse-S pattern (Quiggin, 1982).

5. The intuition for decision under uncertainty

The analysis of uncertainty, presented in this section, is parallel to the analysis of risk. Uncertainty is, however, more interesting because subjective degrees of belief can play a role. Risk is the special case of uncertainty where probabilities are unambiguously known. We briefly describe the uncertainty framework. A set of states (of nature) \( S \) is given. This set is considered to be an exhaustive list of mutually exclusive states: one and only one state will be the true state, but the decision maker is uncertain about which that will be. Subsets of \( S \) are called events. As in Section 1, the outcome set is assumed to be \( \mathbb{R} \). Acts are finite-valued functions from \( S \) to \( \mathbb{R} \). The generic notation of an act is \( (E_1, x_1; \ldots ; E_n, x_n) \). This act yields outcome \( x_j \) if the true state belongs to event \( E_j \). It is implicitly understood in this notation that the events \( (E_1, \ldots , E_m) \) partition the state space.
We assume that the act \((E_1, x_1; \cdots; E_n, x_n)\) is evaluated by the following formula, the **general weighting model**:

\[
\sum_{j=1}^{n} \pi_j U(x_j).
\]

\(U\) denotes the **utility function** and the \(\pi_j\)s are **decision weights**. Decision weights are nonnegative and sum to 1. We assume **monotonicity**, i.e. if for some states of nature the outcomes of an act are replaced by better outcomes then the resulting act is weakly preferred to the original act. This implies that the utility function is nondecreasing. The utility function is assumed to be non-constant so as to avoid triviality. No assumption is yet made about the \(\pi_j\)s and they are permitted to depend on the act in any possible manner. **Subjective expected utility (SEU)** is the special case where the \(\pi_j\)s are **subjective probabilities**, i.e. the following two assumptions hold.

**Assumption 5** [independence of beliefs from tastes]. The decision weight \(\pi_j\) of an event \(E_j\) depends only on the event itself.

With Assumption 5 satisfied, the following assumption can be formulated:

**Assumption 6** [additivity]. The decision weight \(\pi_{AB}\) of a disjoint union \(A \cup B\) is the sum \(\pi_A + \pi_B\) of the decision weights of the separate events \(A\) and \(B\).

There is much interest in relaxations of Assumption 6. First, it is psychologically plausible that people perceive likelihood in a nonlinear manner, a phenomenon which is usually more pronounced under uncertainty than under risk (Currim and Sarin, 1989; Fellner, 1961, p. 684; Tversky and Wakker, 1998; Weber, 1994). A nonlinear processing seems to be as plausible for probabilities as for outcomes, and therefore probability transformation seems to be as useful for descriptive purposes as utility. Second, nonadditive measures of belief, such as Dempster-Shafer belief functions, are extensively used in artificial intelligence (Dempster, 1967; Shafer, 1976). Unfortunately, a relaxation of only Assumption 6 while maintaining full independence of beliefs from tastes turns out to be impossible.

**Theorem 7.** Eq. 4 and Assumption 5 imply subjective expected utility (thus Assumption 6).

Theorem 7 can be interpreted as a negative result. Nonadditive measures cannot be implemented in decisions if Assumption 5 is to be maintained. We therefore turn to a weakening of Assumption 5. The weakening could be interpreted as giving up independence of beliefs from tastes. However, once Assumption 5 is given up, the interpretation of decision weights as indexes of belief, already questionable under expected utility, becomes highly problematic. The interpretation of nonadditive measures, which are simply the decision weights of good- or badnews events, as indexes of belief is, obviously, similarly problematic. Another, more plausible, interpretation of decision weights
is therefore that they are not pure indexes of belief. They may also comprise a component of decision attitude, in addition to the belief component. Under such an interpretation, a decomposition of decision weights into the belief and decision component can be conjectured (Epstein, 1999; Tversky and Wakker, 1998; Wu and Gonzalez, 1999). For consistency with traditional terminology, the name of Assumption 5 is maintained.

A relaxation of Assumption 5 that permits nonadditive measures is provided by Choquet expected utility (CEU), introduced by Schmeidler (1989). His model can be based on the intuition of rank-dependence. That is, the attention paid to an event depends not only on the event but also on how good the outcome yielded by the event is in comparison to the outcomes yielded by the other events. This is the way in which subjective expected utility is generalized.

For the following analysis, consider rank-ordered acts \((E_1, x_1; \cdots; E_n, x_n)\), with \(x_1 > \cdots > x_n\). For event \(E_j\), the ranking position is identified with the event of receiving a worse or equivalent outcome, i.e. it is \(E_j \cup \cdots \cup E_n\). Sarin and Wakker (1998) used the term dominating event for the complement of the ranking position. The following analysis is similar to the analysis under risk. It is presented concisely but in full because it demonstrates the similarity of RDU under risk and CEU under uncertainty, thus the similarity of Quiggin’s (1981) and Schmeidler’s (1989) ideas.

**Assumption 8 [rank-dependence].** The decision weight \(\pi_j\) of an event \(E_j\) depends only on the event and its ranking position.

Next, Choquet expected utility is derived from Assumption 8. The assumption implies in particular that the decision weight of a maximal outcome of a lottery depends only on the associated event \(E\), the ranking position always being the universal event. \(W(E)\) can now be defined as this decision weight. \(W(E)\) is therefore the decision weight generated by the event \(E\) when associated with the best outcome. \(W\) is a capacity, i.e. (1) \(W(\emptyset) = 0\), (2) \(W(S) = 1\), and (3) \(W\) is nondecreasing with respect to set inclusion. (Condition (3) follows from consideration of acts \((A, x; B \cup C, y)\) and \((A \cup B, x; C, y)\) with \(U(x) > U(y)\). Monotonicity implies preference for the first act, which implies that \(W(A \cup B) > W(A)\).)

We express the general weighting model in terms of the capacity \(W\). Consider the act \((E_1, x_1; \cdots; E_n, x_n)\). We assume that the events have been rank-ordered such that \(x_1 > \cdots > x_n\). The decision weight \(\pi_i\) is by definition equal to \(W(E_i)\). Next consider a general \(i\).

**Observation.** The total decision weight assigned to outcomes \(x_1, \ldots, x_i\), i.e. \(\pi_1 + \cdots + \pi_i\), is \(W(E_1 \cup \cdots \cup E_i)\).

**Explanation.** Consider the acts \((E_1, x_1; \cdots; E_i, x_i; E_{i+1}, x_{i+1}; \cdots; E_n, x_n)\) and \(((E_1 \cup \cdots \cup E_i), z; E_{i+1}, x_{i+1}; \cdots; E_n, x_n)\) for any outcome \(z\) exceeding \(x_{i+1}\), e.g., \(z = x_i\). Because decision weights must sum to one, \(\pi_1 + \cdots + \pi_i = 1 - \pi_{i+1} = \cdots = \pi_n = W(E_1 \cup \cdots \cup E_i)\). Note that, by Assumption 8, the outcomes \(x_{i+1}, \ldots, x_n\) all have the same ranking position in the two acts and therefore the same decision weights.
The observation implies that the decision weight \( \pi_i \) of event \( E_i \) is the difference \( W(E_i \cup \cdots \cup E_j) - W(E_i \cup \cdots \cup E_{j-1}) \). It is standard that this difference is \( \pi_i = W(E_i) \) for \( i = 1 \).

The rank-ordering of the events was crucial in our analysis. Let us summarize and give the formal definition of \textit{Choquet expected utility (CEU)}. For \( x_1 > \cdots > x_n \),

\[
CEU(E_1, x_1; \ldots; E_n, x_n) = \sum_{j=1}^{n} \pi_j U(x_j)
\]

where

\[
\pi_j = W(E_1 \cup \cdots \cup E_j) - W(E_1 \cup \cdots \cup E_{j-1}).
\]

**Conclusion 9.** Under the general weighting model (Eq. 4), monotonicity and Assumption 8 imply CEU.

Empirical measurements of decision weights have been described by Fox, Rogers, and Tversky (1996), Fox and Tversky (1995), Kilka and Weber (1999), and Wu and Gonzalez (1999). We hope that the preceding explanation clarifies that the intuitive basis of CEU is the same as of RDU. Thus, a psychological background has also resulted for Schmeidler’s (1989) Choquet expected utility. It will be argued in the next section that, given this intuition, the comonotonicity condition is not just a mathematical tool but is a natural concept. Let us now turn to a discussion of pessimism.

Pessimism means again that the attention paid to an event gets higher as the event gets rank-ordered worse. That is, assume that event \( E \) yields outcome \( x \) and \( D \) is the ranking position of \( E \). Then the decision weight of \( E \) is \( W(D^c \cup E) - W(D^c) \). Under pessimism, worsening the ranking position (decreasing the event \( D \) of receiving something worse) increases the decision weight of \( E \). That is, if \( C \subseteq D \), then

\[
W(C^c \cup E) - W(C^c) \geq W(D^c \cup E) - W(D^c).
\]

Similar to risk, a capacity \( W \) satisfying Eq. 6 is called \textit{convex}. Eq. 6 can be rewritten as \( W(A \cup B) + W(A \cap B) \geq W(A) + W(B) \) after appropriate substitution of symbols (left to the reader). Optimism is similarly characterized by \textit{concavity} of the capacity, i.e. Eq. 6 with \( \leq \) instead of \( \geq \).

6. Coalescing and comonotonicity

Both in risk and in uncertainty, the rank-dependent formulas have been given for distinct outcomes \( x_1 > \cdots > x_n \). Eqs. 3 and 5 can also be used if the inequalities are weak, i.e. \( x_1 \geq \cdots \geq x_n \). These claims follow from substitution and are left to the reader. For an act \((E_1, x_1; \ldots; E_n, x_n)\) with \( x_1 = x_{n+1} \), the decision weight and the ranking position of event \( E_j \) depend on the chosen rank-ordering between \( x_i \) and \( x_{i+1} \). This choice can be made arbitrarily and is immaterial for the associated preferences.
We next discuss comonotonicity, introduced by Schmeidler (1989). The condition has sometimes been criticized. An explanation of its intuition therefore seems to be worthwhile. For simplicity, assume a finite state space $S = \{s_1, \ldots, s_n\}$. For a permutation $(\rho_1, \ldots, \rho_n)$ of $(1, \ldots, n)$, consider the set $C^\rho = \{f \in \mathbb{R}^n : f_{\rho_1} \geq \cdots \geq f_{\rho_n}\}$. It can be seen that $C^\rho$ is a convex cone. For all acts in the cone $C^\rho$, we can use the same decision weights $\pi_{\rho_j}$ determined by

$$\pi_{\rho_j} := W(s_{\rho_1}, \ldots, s_{\rho_j}) - W(s_{\rho_1}, \ldots, s_{\rho_{j-1}}, s_{\rho_{j+1}}, \ldots, s_{\rho_n})$$

in the computation of CEU. If acts are in the same cone, then $f_j > f_j$ and $g_j > g_j$ for no states $s_j$ and $s_j$. Acts belonging to the same cone are called comonotonic.

Within comonotonic sets, CEU coincides with an SEU functional. This SEU functional is defined by taking the CEU utility function and taking as probabilities the decision weights $\pi_{\rho_j}$ associated with the comonotonic set. Therefore, CEU exhibits many characteristics of SEU within comonotonic sets. In particular, it satisfies the same preference axioms.

The comonotonic agreement of CEU with SEU is implied by the theory but is also empirically interesting. Consider acts belonging to different comonotonic sets. The states of nature are rank-ordered differently for such acts. This difference will enhance variations in the psychological attention paid to the states. Subjects will exhibit more pronounced violations of SEU, due to pessimism, optimism, etc. When only acts are considered from one comonotonic set, fewer violations of SEU can be expected. According to CEU theory, the effects of pessimism and optimism will then be kept constant. In reality, they can be expected to be smaller than when the rank-ordering of the acts varies.

Comonotonicity is extensively used in preference axiomatizations of CEU. Most axiomatizations consist of restricting the SEU axioms to comonotonic acts. For a continuum of outcomes, CEU holds as soon as SEU holds within every comonotonic set (this is easily derived from Wakker and Tversky, 1993, Proposition 8.2). An empirical application of comonotonicity can be found in utility measurement. Wakker and Deneffe (1996) demonstrated that utility can be measured under CEU by restricting SEU techniques to comonotonic sets of acts. Such a restriction has the empirical advantage of avoiding the biases generated by rank-dependence, and therefore seems desirable.

Some authors have pointed out that rank-dependence and comonotonicity are often used as technical tools and that there is still a need for an intuitive foundation (Luce, 1996a, p. 85; Luce, 1996b, p. 304; Safra and Segal, 1998, p. 28). Our paper has provided such a foundation, building on ideas described before in the literature. We have argued that rank-dependence and comonotonicity do have intuitive and empirical merit. Yaari (1987, p. 104) already emphasized the intuitive importance of comonotonicity when discussing his central axiom (dual independence): “The foregoing proposition makes it clear that the economic interpretation of dual independence lies in the intuitive meaning of comonotonicity.”

Obviously, alternative derivations of CEU and RDU that do not use rank-dependence or comonotonicity in their axioms are also interesting. Such derivations are given by Luce
(1998) and Safra and Segal (1998). In these derivations, rank-dependence follows from other conditions.

7. Conclusion

This paper has argued that RDU is not just a mathematical device but that it is based on intuition and has real-world merits. The intuition of rank-dependence was described in terms of decision weights. The RDU formula naturally followed as well as empirically meaningful preference conditions. Optimism and pessimism were explained in terms of the intuitive foundation. An analogous reasoning was applied to the uncertainty case and a psychological background for Schmeidler's (1989) Choquet expected utility resulted. Once the intuition is understood, comonotonicity conditions and rank-dependence are no longer mere theoretical tools. They become natural concepts upon which preference conditions, empirical tests, and improvements of utility measurement can be based.

Our preference for RDU, and we believe also its general popularity, are based not only on its mathematical or empirical performance but also on an intuitive aspect of the model: The nonlinear weighting of chance, and nonadditive measures of belief, have the potential of becoming useful concepts, not only in economics but also in other areas such as psychology and artificial intelligence.

Appendix A. Related literature on the intuition of rank-dependence

This appendix discusses some intuitive arguments for rank-dependence that have been presented in the literature. A first example from the psychological literature is Birnbaum's (1974) study of the formation of personality impressions. For example, Birnbaum studied the likableness of a person on the basis of intellectuality, shyness, loyalty, etc. He found empirical violations of additive aggregation and proposed a configural weighting model that better describes how intellectuality etc. are aggregated into likableness of a person. Under configural weighting, "... the weight of a stimulus depends upon its rank within the set to be judged" (p. 559). Although this model is formally different from RDU, it does already contain an intuition of rank-dependence. Configural weighting theory was later extended to risky choices (Birnbaum and Navarrete, 1998 and the references therein).

A remarkable study is Lopes (1984) who argued for the intuitive value of rank-dependence in risk theory as an extension of the Gini index of inequality. The rank-dependent aspect of such measures of inequality was formulated by her as "... embody distributional objectives in terms of the relative weight given to inequality at different points on the income scale ... The central psychological premise in this article is that people's intuitions about risks are functionally similar to intuitions about distributional inequality. ... representation that captures psychologically salient features of risky distributions" (p. 468). She then explained that people, well aware of the objective probabilities, still "may wish to weight outcomes differently at different points in the distribution" and discusses human ways of reasoning reflecting this procedure (p. 469). Experiments were
presented to test for the role of rank-dependence. Lopes concluded that rank-dependence (called the distributional model) "... seems to offer the potential of capturing in a psychologically meaningful way many interesting and important features of people's processing of and preference for risks" (p. 484).

Let us emphasize that Lopes (1984) derived her ideas solely from intuition and psychological principles. No preference axioms were considered. Her work was developed independently of Quiggin (1981, 1982) or other presentations of RDU. Lopes (1987) presented experiments where subjects were asked to speak aloud about their motives for choices between multiple outcomes gambles. It turned out that subjects paid much attention to goodnews events (receiving at least as much as ...) and, similarly, badnews events. This attention is formalized through the probability weighting function in rank-dependent theories. Rank-dependent decision weights then result from difference-taking. The great attention to good- and badnews events also supports the inverse-S shapes of the weighting functions.

A third example from the psychological literature is Weber (1994). She used a somewhat different approach than this paper, invoking an analogy with estimation theory and asymmetric loss functions, and concluded "these processes need not necessarily be perceptual in origin. Instead, in this article, I argued that configural or rank-dependent weighting could be interpreted as strategic or motivational (i.e. a reasonable response that takes into consideration existing constraints that are ignored by the expected utility model)" (p. 236). On p. 237 she discussed perceptual origins ("attentional salience"): "... and more extreme outcomes may get greater weight than outcomes in the middle of the distribution, simply because they are more noticeable."

Models that pay special attention to highest or lowest outcomes can be considered to be special cases of rank-dependence. An example is Rawls' (1971) proposal for welfare evaluation, where all importance weight is allocated to the poorest person in the society. Rank-dependent models for welfare were developed by Ebert (1988) and Weymark (1981). Similar models, with the importance weight divided over the highest and lowest outcomes, were proposed by Arrow and Hurwicz (1972) and Hurwicz (1951). Models that deviate from expected utility only by overweighting highest and/or lowest outcomes were proposed by Bell (1985), Cohen (1992), Gilboa (1988), and Jaffray (1988). In time preference, rank-dependence arises when people are especially sensitive to decreases in salary. This is a special case of rank-dependence, related to the immediately preceding period (Gilboa, 1989; Shalev, 1997).

Yaari (1987) related the intuitive meaning of comonotonicity to hedging. Hedging concerns combinations of outcomes and therefore requires a linear structure on the outcome set. For example, consider two gambles for money on the same toss of a coin. The first gamble is \((H, 30; T, 10)\) yielding $30 if heads comes up and $10 if tails comes up. The second gamble is \((H, 10; T, 30)\). These two gambles are equivalent but their fifty-fifty-outcomes-mixture \((H, 20; T, 20)\) is usually preferred. In the mixture, a reduction of risk has resulted. Hedging occurs because the good outcome of one lottery neutralizes the bad outcome of the other lottery and vice versa. The lotteries serve as complementary goods. The described neutralization can occur only because the gambles are not comonotonic. Hence, Yaari argued that an independence condition (his Axiom A5) is natural only in the...
absence of hedging, i.e. only for comonotonic gambles. A similar argument was presented by Roell (1987). Hedging is central in the portfolio selection of assets.

Schmeidler (1989) used a similar framework that generalizes Yaari's model in two respects. First, Yaari considered real-valued outcomes (interpreted as money), whereas Schmeidler dealt with general convex outcome sets (interpreted as probability distributions over prizes). Second, Schmeidler assumed states of nature for which no probabilities need to be given. Yaari's model can be considered the special case of Schmeidler's model where probabilities of the states of nature are given and outcomes are one-dimensional.

We next discuss Quiggin (1982). He first discussed what he called the primitive approach, i.e. our Eq. 2, transforming only probabilities of fixed outcomes, and pointed out: "the fundamental problem in these theories is that any two outcomes with the same probability must have the same decision weight. This fails to take account of the fact that while individuals may distort the probability of an extreme outcome in some way, they need not treat intermediate outcomes with the same probability in the same fashion." A similar intuition can be recognized in Fellner (1961, p. 674/675). So as to formalize this observation, Quiggin proposed to order the possible outcomes \( x_i \) and the corresponding probabilities \( p_i \) in each prospect and denoted the rank-ordered probability vector \( (p_1, \ldots, p_n) \) by \( p \). Quoting again from his paper: "The anticipated utility function is defined to be \( V = h(p)U(x) = \sum_i h_i(p)U(x_i) \) where \( U \) is a utility function with properties similar to that of von Neumann-Morgenstem, while \( h(p) \) is a vector of decision weights satisfying \( \sum_i h_i(p) = 1 \). In general, \( h_i(p) \) depends on all the \( p_i \)'s and not just on \( p_i \). Thus, for example, the fact that \( p_j = p_k \) would not imply that \( h_j(p) = h_k(p) \)."

Quiggin's formula is a special case of the general weighting model where the decision weights are independent of the outcomes given the rank-ordered probability vector \( p \). Quiggin gave preference conditions to characterize his formula. He then showed that RDU follows from a continuity condition. The purpose of our analysis was different. We did not invoke technical conditions such as continuity in the derivation but derived rank-dependence from intuitive arguments.

Appendix B. Proofs

**Proof of Theorem 2.** Decision weights always sum to one and, hence, \( w(p_1 + p_2) = 1 - w(1 - p_1 - p_2) = w(p_1) + w(p_2) \). Consequently, \( w \) satisfies Cauchy's equation. By Aczél (1966), \( w \) must be linear. Note here that \( w \) is bounded by 0 and 1 so that no nonlinear solutions to the Cauchy equation are possible. \( w \) is the identity function because \( w(0) = 0 \) and \( w(1) = 1 \).

**Proof of Theorem 7.** For each event \( E \), define \( W(E) \) as the decision weight of that event. \( W(E) \) is nonnegative, \( W(\emptyset) = 0 \), and \( W(S) = 1 \). Decision weights of partitions always sum to one, therefore \( W(E_1 \cup E_2) = 1 - W((E_1 \cup E_2)^c) = W(E_1) + W(E_2) \). \( W \) is a probability measure and SEU follows.
Acknowledgments

This paper received helpful comments from Han Bleichrodt and Paolo Ghirardato.

Notes

1. This dependence is not expressed in the notation.

References


