TRIPLE PREFERENCE THEORY OF CHOICE UNDER RISK AND UNCERTAINTY

Janne Gustafsson

Cheyne Capital Management Limited
Stornoway House, 13 Cleveland Row, London SW1A 1DH, United Kingdom
Tel. +44 (0) 20 7031 7569
Fax +44 (0) 20 7031 7669
Email janne.gustafsson@cheynecapital.com

April 2006
Abstract: This paper develops a theory of choice under risk and uncertainty which, together with the associated definition of risk aversion, separates the concepts of risk aversion and diminishing marginal utility from each other. Building on separate definition of the decision maker’s two elementary preference relations—preferences over lotteries and preferences over consequences—and a set of consistency properties that link these two together, the paper derives a third preference relation on a set of “utility lotteries,” and shows that the preference functional for lotteries necessarily separates into the representation functions of the two other preference relations of the theory. The paper then demonstrates that a reasonable definition of risk aversion necessitates the use of cardinal utility, whereby it is also possible to achieve the separation result of the paper. The theory is compatible with expected utility theory, and allows for multi-attribute consequences, mean-risk preferences and state-dependent utility.

Keywords: Decision theory; Choice under risk and uncertainty; Risk attitudes

JEL Classification: C44, D81

Acknowledgements: The author would like to thank the seminar participants at the McCombs School of Business at the University of Texas at Austin, and at the Fuqua School of Business at the Duke University. The author is also grateful to Robert T. Clemen, James S. Dyer, Lorenzo Garlappi, Harry M. Markowitz, Luca Rigotti, Ahti Salo, and James E. Smith for comments and suggestions.
1 INTRODUCTION

Expected utility theory of von Neumann and Morgenstern (1947) benefits from several appealing characteristics as a theory of a decision maker’s (DM) behavior under risk: It has an elegant axiomatic foundation and it leads to a lucid model of expected utility. This notwithstanding, the expected utility model has been subjected to several criticisms. Most of these concern the independence axiom (see, e.g., Samuelson (1952)), which underlies the use of expectation (i.e., linearity in the probabilities) in expected utility theory. As a result, a wide array of alternative theories of decision under risk have been developed: prospect theory of Kahneman and Tversky (1979), the rank-dependent model of Quiggin (1982, 1993), Machina’s (1982) analysis of Fréchet differentiable preference functionals, and dual theory of Yaari (1987), among others (for a review, see Sugden (1986), Machina (1987) and Starmer (2000)).

This paper presents a theory of choice under risk and uncertainty, which is primarily motivated by another kind of critique put forth by Allais (1953, 1979a, 1979b). Allais argues that the DM should consider the entire form (all moments) of the utility distribution (a probability distribution defined over a utility scale) rather than its expectation only. He bases this argument on the reasoning that the concept of risk must be related to the dispersion of utility rather than to the dispersion of the actual outcomes of the lottery, because according to the classic notion of utility (see Jevons (1871/1970)), the DM always considers the utility of an outcome rather than its actual value (e.g., monetary value or quantity; outcomes can also be non-numeric) and hence he/she also does so when analyzing a lottery of outcomes. Since this observation conflicts with the results of expected utility theory, which is concerned with the first moment of a utility distribution only, Allais (1979a, p. 104) concludes that this theory must be erroneous and writes that “the fundamental error of the entire American school is its indirect and unconscious neglect of the dispersion of psychological values.”

The proponents of expected utility theory, however, have rejected Allais’s claim by
asserting that the consideration of the dispersion of utility would result in double-counting of the DM’s risk attitude, since the curvature of the utility function accounts for risk already once (see, e.g., Amihud (1979, p. 155)). Still, if we accept Allais’s standpoint concerning the nature of risk, it is quite difficult to see how a utility function could capture the dispersion of a probability distribution that is created from the actual outcomes of a lottery by using the utility function itself, unless the concept of utility in expected utility theory is fundamentally different from what Allais means by utility (i.e., classic non-stochastic cardinal utility). If so, a question arises as to how these two concepts of utility are related to each other and whether the von Neumann-Morgenstern utility function accounts for the dispersion of classic cardinal utilities, which Allais calls for.

Even if Allais were wrong about the nature of risk aversion and it were the dispersion of monetary values that should be accounted for, the variance of monetary values, as well known, can be captured through a von Neumann-Morgenstern utility function only when additional restrictions are made (e.g., the use of a quadratic utility function or normal distributions only; Markowitz (1959, 1987)), which raises the question whether the consideration of the dispersion of monetary values and the use of a von Neumann-Morgenstern utility function are actually the same thing. Indeed, as Yaari (1987) puts it, at the level of fundamental principles risk aversion and diminishing marginal utility are “horses of different colors.” A risk averse DM seeks to avoid the possibility of getting a poor outcome while one can have diminishing marginal utility only because a loss of one unit of a commodity hurts a rich man less than a poor man.

Motivated by Allais’s and Yaari’s observations, the primary purpose of this paper is to separate the concepts of diminishing marginal utility and risk aversion from each other in the theory of choice under risk and uncertainty. However, conventional preference modeling approaches do not make it possible to distinguish risk preferences from among preferences over lotteries, mainly because they do not formalize which part of the DM’s preferences over lotteries is purely related to consequences, thereby inseparably confounding the DM’s preferences over consequences and preferences for risk together.
Thus, the accomplishment of the primary purpose of this paper necessitates the development of a theory of choice under risk and uncertainty that separately formalizes the DM’s preferences over consequences and preferences over lotteries, and links these two together through necessary consistency properties (for example, preference order on lotteries always yielding the same sure outcome must be the same as preference ordering of the sure outcomes).

Indeed, a distinctive feature of the present axiomatic system is that it relies on the definition of preference relations on three distinct sets of elements: (i) the DM’s preferences over consequences, (ii) his/her preferences over lotteries, and (iii) his/her preferences over “utility lotteries,” which are lotteries where each outcome is replaced by the utility of that outcome. Hence, the present development is called triple preference theory. The last preference relation is motivated by the axioms of the theory and is conceptually needed to formalize a preference relation that represents everything about the DM’s preferences over lotteries that the DM’s preferences over the lotteries’ outcomes (i.e. consequences) do not represent. In a broad sense, when probabilities of lotteries’ outcomes are known, these preferences can be called the DM’s risk preferences. (In the presence of ambiguity, preferences over utility lotteries will capture both the DM’s risk attitude and ambiguity attitude.) In particular, this paper shows that, under very mild conditions, the DM’s preference model for lotteries necessarily separates into two components: (i) a certainty equivalent operator, which can be a functional like the expectation operator or a mean-risk model, and (ii) a utility / value function, where the former is the representation of DM’s preferences over utility lotteries and the latter the representation of the DM’s preferences over consequences. The paper then demonstrates that this preference model, when coupled with a revised definition of risk aversion which accounts only for the part of the DM’s preferences that can reasonably be assumed to be related to risk / dispersion, leads to the separation of the concepts of diminishing marginal utility and risk aversion.

It is imperative to highlight that the present paper does not aim at developing a theory of choice under risk and uncertainty that would compete with other normative theories of
choice, such as expected utility theory. Rather, the present theory is constructed so that the separation result can be used in conjunction with several earlier theories of choice under risk and uncertainty. A natural pair to triple preference theory is Savage’s (1954) subjective expected utility theory, the axioms of which can easily be added on the top of the axioms of the present system in order to set additional rationality properties for the DM’s preferences over lotteries. Indeed, when the DM agrees with the axioms of triple preference theory and subjective expected utility theory, it can be shown that the DM’s risk attitude is related to the dispersion of classic cardinal utility, and at the same time, there exists a subjective expected utility preference functional representing the DM’s preferences over lotteries, where the utility function can be different from the classic cardinal utility function used to assess the DM’s risk attitude. In particular, it turns out that the DM’s risk attitude is related to the transformation between the von Neumann-Morgenstern utility function and the classic cardinal utility function, which, interestingly enough, coincides with the postulate in Dyer and Sarin’s (1982) paper on relative risk aversion.

The remainder of the paper is structured as follows. Section 2 reviews briefly some central approaches in decision theory. Section 3 presents the axioms of the present framework and develops the general separable preference model. Section 4 discusses some certainty equivalent operators and the relationship of the present framework to expected utility theory. Section 5 examines the definition of risk attitudes and presents the final separation result of the paper. Section 6 summarizes the main results and draws conclusions.

2 EARLIER APPROACHES

Broadly speaking, normative analysis of decision making can be separated into (i) choice under certainty, (ii) choice under risk, and (iii) choice under uncertainty. As defined by Knight (1921), choice under risk refers to a setting where objective probabilities exist, while choice under uncertainty is used to mean a setting where objective probabilities cannot be assigned to uncertain events.
2.1 **Choice under Certainty**

Choice under certainty has extensively been studied by economists (see e.g. Jevons (1871/1970) and Krantz et al. (1971)) and researchers of operations research and decision analysis (see e.g. Keeney and Raiffa (1976), Dyer and Sarin (1979) French (1986), Clemen (1996)). These two lines of research have somewhat differing terminology for the same mathematical objects. In particular, in order to make a distinction to von Neumann-Morgenstern utility functions, decision analysts refer to ordinal utility functions as ordinal value functions (compare Mas-Colell et al. (1995) to French (1986) and Clemen (1996)) and to classic cardinal utility functions as measurable value functions (compare e.g. Jevons (1871/1970) to Dyer and Sarin (1979); see also Krantz et al. (1971)). Also some economists, such as Kahneman and Tversky (1979), use the value function terminology to highlight that they are not employing von Neumann-Morgenstern utility functions. Nevertheless, in the economic sense, all these functions can be called utility functions, as they all measure the desirability of consequences to the DM.

In general, we can make a distinction between three classes of utility functions: (i) ordinal utility / value functions, (ii) classic non-stochastic cardinal utility functions (i.e. measurable value functions), and (iii) von Neumann-Morgenstern utility functions. The last two classes of utility functions can both be regarded as *cardinal* in the sense that they can only be subjected to positive affine transformations if they are to preserve a desirable property. The classic cardinal utility function will preserve the property of representing the DM’s preferences over differences of consequences so that after being subjected to a positive affine transformation the utility function still correctly measures diminishing marginal utility from additional units of commodities. For example, if under one utility function the utility difference between owning 2 cars and 1 car is smaller than the difference between owning 1 car and 0 cars, then the same also holds under all positive affine transformations of the utility function. Similarly, only a positive affine transformation of a von Neumann-Morgenstern utility function will preserve the expected utility representation of the DM’s preferences over lotteries. Classic cardinal utility functions and von Neumann-Morgenstern utility functions are, in general, different (for discussion, see Ellsberg (1954) and Baumol (1958)), because the latter involves the
element of risk, as it is elicited under risk by comparing lotteries, but the former does not, as it can be elicited completely under certainty by comparing the desirability of differences of consequences. Nevertheless, we know that these two classes of cardinal utility functions must be related to each other through a strictly increasing transformation, because they both belong to the same set of admissible ordinal utility functions (see e.g. Ellsberg (1954)).

2.2 CHOICE UNDER RISK AND UNCERTAINTY

Despite the development of over a dozen non-expected utility theories in the past three decades, expected utility theory (EUT) of von Neumann and Morgenstern (1947) remains as the single generally accepted normative theory for decision making under risk. This can be mainly attributed to one unique property of EUT: Among theories relying on preference functionals, it is the only one that leads to a *dynamically consistent* preference model (Machina (1989)). That is, in a dynamic setting an expected utility maximizer always sticks to his/her pre-planned contingency strategy as time passes by and uncertainties resolve, which is not the case with decision makers in non-expected utility theories.

There are four main axiomatic systems leading to the expected utility model. First, there is the standard axiomatic system of EUT, which appears e.g. in Fishburn (1970) and involves the *independence* axiom. Second, there is Luce and Raiffa’s (1957) system that relies on the analysis of compound lotteries and on the *substitutability* axiom (see also French (1986) for a variant of this system). Third, one of the landmarks in expected utility theory is Savage’s (1954) *subjective expected utility theory* (SEUT), where the central axiom is called the *sure-thing principle*, which is similar to the independence axiom in its essence. The special property of this theory is that it does not require the existence of (objective) probabilities for lotteries’ outcomes, but rather implies that there must exist “subjective probabilities” that, together with the DM’s utility function, represent the DM’s preferences over lotteries. Finally, there is the Anscombe-Aumann (1963) system, which is analogous to Luce and Raiffa’s (1957) development in terms of relying on the assumptions of substitutability and reduction of compound lotteries but
differs in that it is applicable in the subjective probability context.

Indeed, the key difference between these formulations is that the two first systems are based on the existence of objective probabilities, and hence are applicable only in choice under risk, while the two latter ones are specifically applicable when it is impossible to attach probabilities with the likelihood of lotteries’ outcomes; that is, these systems are applicable in choice under uncertainty. In particular, the two latter developments show that when plausible axioms about the DM’s preferences over acts (lotteries without probabilities) hold, the DM’s preferences model for acts can be represented by an expected utility model so that there exists some (“subjective”) probability measure and some utility function that together represent the DM’s preferences over lotteries. Triple preference theory is very much like Savage’s (1954) theory in the sense that it shows that when certain plausible axioms about the DM’s preferences over lotteries (acts) hold, the DM’s preference model for lotteries will necessarily separate into two components.

Several non-expected utility theories of choice under uncertainty have also been developed in response to the observed violations of EUT and SEUT in practice. These include, for example, the Choquet-expected utility theory of Schmeidler (1989) who shows that under certain axioms, there exists a possibly non-additive capacity measure (instead of an additive probability measure) and a von Neumann-Morgenstern utility function that together represent the DM’s preferences under those axioms. For a review of non-expected utility models for choice under uncertainty, we refer to Camerer and Weber (1992). Still, similar to their cousins usable in choice under risk, these theories have enjoyed limited use both in theory and practice.
3 THEORY

3.1 OVERVIEW

In order to clarify the nature of the present development, it is instructive to review some basic facts about triple preference theory. At the general level, triple preference theory is a generic theory of choice under risk and uncertainty with the aim of separating the concepts of risk aversion and diminishing marginal utility from each other. As such, the theory imposes minimal rationality properties on the DM’s preferences over lotteries and preferences over consequences: For example, it does not assume the existence of a probability measure or assume any structure for consequences. The theory can, therefore, be immediately applied both in choice under uncertainty and in multi-attribute decision making. It also allows for state-dependent utility and mean-risk-type preferences, whereby it also covers the mean-risk objective functions used in my earlier paper on project portfolio selection (Gustafsson and Salo (2005)). It is also worth highlighting that the theory assumes only the existence of an ordinal utility function, which one might consider a curious property for a theory of choice under risk, as conventional theories almost invariably assume the existence of a cardinal utility function. However, it turns out that there is no particular need to introduce cardinal utility, except when it comes to the definition of risk aversion and possibly to the elicitation of the certainty equivalent operator.

Triple preference theory starts with the definition of the two elementary preference relations of the DM, (i) preferences over consequences (sure outcomes) and (ii) preferences over lotteries (acts), followed by the specification of consistency requirements that link these two relations together. Notably, these consistency requirements motivate the introduction of a third preference relation which is defined on the set of utility lotteries. This third preference relation is later employed in the derivation of the separation result of the paper.
Triple preference theory contains a total of eight axioms. Axioms 1 and 2 define a weak ordering on the set of consequences and ensure that there is an ordinal utility function representing this ordering. This part of the theory is a completely standard development of ordinal utility functions in economic theory. Axioms 3 and 4 define a weak ordering on the set of lotteries and provide a necessary condition required to ensure the existence of a preference functional representing the DM’s preferences over lotteries. Axiom 5, called substitutability of consequences, is the first of the two axioms linking the DM’s preferences over consequences to his/her preferences over lotteries, and the primary element motivating the introduction of the concept of a utility lottery. Axiom 5 states that whenever (a) two lotteries are identical except in terms of their outcome in a single state of nature and (b) the DM is indifferent between those differing outcomes, then the DM should be indifferent between the lotteries as well. Axiom 6 and 7 define a weak ordering on the set of utility lotteries and link this preference relation to the preferences over lotteries through a necessary consistency requirement called indifference class relationship. Finally, Axiom 8, called consistency with strict absolute dominance, is the second link between the DM’s preferences over consequences and preferences over lotteries. The axiom states that if the outcome of a lottery is at least as good as the outcome of another lottery in each state of nature, and the outcome of the first lottery is strictly better than that of the second lottery at least in one state, then the first lottery must be preferred to the second. Thus, Axiom 8 is essentially the monotonicity axiom of the system.

3.2 LOTTERIES

Let $\mathcal{L}$ be the set of all relevant lotteries under consideration. The outcomes of these lotteries can be either consequences (sure outcomes) in a set $C$ or other lotteries. All lotteries are finitely compounded. For the purposes of this paper, the nature of consequences in $C$ needs not be defined in more precision; however, they can be, for example, multi-attribute alternatives involving $n$ attributes (mathematically, such consequences would be $n$-tuples).

To simplify our development, we make the usual reduction of compound lotteries
assumption. Therefore, we can limit our analysis to all simple lotteries that result from lotteries in $\mathcal{L}$. Let us denote the set of these simple lotteries by $L$. Simple lotteries are modeled as functions from state space $S$ to consequence space $C$, $\tilde{x}: S \rightarrow C$, called acts. When there is also a probability space $(S, \mathcal{F}, \mathbb{P})$ defined over $S$ (where $\mathcal{F}$ is the $\sigma$-algebra on $S$, and $\mathbb{P}$ the associated probability measure), simple lotteries are, by definition, $C$-valued random variables. For the sake of simplicity, we will use the terms “simple lottery” and “lottery” interchangeably in the following; this can be done without risk of confusion, because the set $\mathcal{L}$ is not further used in the ensuing development.

**Definition 1: Set of lotteries.** Set of lotteries is denoted by $\mathcal{L}$.

**Definition 2: Set of simple lotteries.** Set of simple lotteries is denoted by $L$.

**Definition 3: Set of states.** Set of states of nature is denoted by $S$.

**Definition 4: Set of consequences.** Set of consequences is denoted by $C$.

### 3.3 Elementary Preferences

#### 3.3.1 Preferences over Consequences

Let us now define the DM’s preferences over outcomes of lotteries. This is a completely standard development of ordinal utility functions. Let there be a weak order (transitive and complete) $\succeq_C$ on $C$ representing the DM’s preferences over consequences. Let us also assume that there is a countable order dense subset of $C$, which ensures that there exists an ordinal utility function $u(.)$ on $C$ (Krantz et al. (1971), French (1986)). This function is unique up to a strictly increasing transformation, i.e., if $u(.)$ is an ordinal utility function and $\phi(.)$ is a strictly increasing function, then also $\nu(.) = \phi(u(.))$ is an ordinal utility function. These assumptions can be formalized as follows.

**Axiom 1: Weak ordering of consequences.** The DM’s preferences over $C$ form a weak order $\succeq_C$.

**Axiom 2: Order dense subset.** There is a countable (finite or countably infinite) order dense subset of $C$. That is, there is a countable subset $B$ of $C$ so that for all $x$ and $z$ in $C$, such that $x \succeq_C z$, there exists $y$ in $B$ such that $x \succeq_C y \succeq_C z$. 

10
**PROPOSITION 1: Ordinal utility function.** There is a utility function $u$ that represents $\succ_{C}$ in the sense that $x \succ_{C} y \iff u(x) \geq u(y) \forall x, y \in C$. Moreover, $u$ is ordinal in the sense that also $v(x) = \phi(u(x))$, where $\phi$ is a strictly increasing function, is a representation of $\succ_{C}$.

**PROOF:** Follows from Axioms 1 and 2. See Krantz et al. (1971).

### 3.3.2 Preferences over Lotteries

Let us next define the DM’s preferences over lotteries under consideration. Let there be a weak order $\succeq_{L}$ on $L$ representing the DM’s preferences over lotteries. To ensure the existence of a real-valued preference functional, we assume that for each lottery $\tilde{x}$ in $L$, there exists an equally preferable degenerate lottery $\tilde{\delta}$, which is an act that has only a single possible outcome $c$. This outcome is called the (consequence) certainty equivalent of $\tilde{x}$. Mathematically, $\tilde{\delta}:S \to C, \tilde{\delta}(s) \equiv c$, is a constant function from $S$ to $C$. This existence assumption serves the purpose of EUT’s continuity and reference experiment axioms by stating that the DM is always able to find a sure outcome that is equally preferable to a given lottery and that the related degenerate lottery belongs to $L$.

**AXIOM 3: Weak ordering of lotteries.** The DM’s preferences over $L$ form a weak order $\succeq_{L}$.

**AXIOM 4: Existence of an equally preferable degenerate lottery.** For each $\tilde{x} \in L$, there is a degenerate lottery $\tilde{\delta} \in L$ such that $\tilde{\delta} \sim_{L} \tilde{x}$.

### 3.4 Substitutability of Consequences

The substitutability of consequences axiom is a weaker version of EUT’s substitutability axiom. Instead of allowing the replacement of both sure and risky outcomes with equally preferable sure or risky outcomes, the axiom is limited to sure outcomes only. The axiom is motivated by the logic that if the DM is truly indifferent between two consequences, then he/she should not care which one of the consequences he/she receives as an outcome of a lottery.
For example, suppose that the DM is indifferent between two consequences, say, an apple and an orange. Then, it sounds reasonable to assume that the DM should be indifferent between the two following lotteries:

For monetary lotteries, the axiom may seem redundant, because under typical preferences for money there do not exist two different amounts of money that the DM would hold equally preferable. However, we will observe that the substitutability of consequences axiom plays a major role in the separation of the DM’s preferences into two components. The axiom is formalized as follows.

**AXIOM 5: Substitutability of consequences.** Let \( b \) and \( c \) in \( C \) be such that the DM holds \( b \sim_c c \). Let \( \tilde{x} \) be a lottery in \( L \) which yields \( b \) in state \( s^* \), i.e. \( \tilde{x}(s^*) = b \). Then, a lottery \( \tilde{y} \), which yields the same outcomes as \( \tilde{x} \) in each state of nature except in state \( s^* \) where it yields \( c \), is equally preferable to \( \tilde{x} \), i.e. \( \tilde{x} \sim_L \tilde{y} \).

It needs to be highlighted that by removing the possibility to replace risky lotteries with equally preferable sure outcomes, this version of the substitutability axiom also loses the property that makes it so central in EUT. Therefore, the implications of the axiom are not as far-reaching as what they would be if the usual strong substitutability would hold. However, it is important to realize that the strong substitutability of EUT also implies this weak version of substitutability, and that, therefore, rejection of Axiom 5 would also necessarily lead to rejection of EUT, too.
The main consequence of the substitutability of consequences axiom is that all lotteries that yield the same amount of utility in each state of nature must be equally preferable. We formalize this claim as the following proposition.

**PROPOSITION 2: Equal preference of lotteries with equally preferable outcomes.** If lotteries \( \tilde{x} \) and \( \tilde{y} \) yield equally preferable outcomes in each state of nature, they are equally preferable. That is, for all \( \tilde{x} \) and \( \tilde{y} \) in \( L \), \( \tilde{x}(s) \sim_c \tilde{y}(s) \forall s \in S \implies \tilde{x}_L \sim \tilde{y}_L \).

**PROOF:** Let \( \tilde{x} \) and \( \tilde{y} \) be any lotteries for which it holds \( \tilde{x}(s) \sim_c \tilde{y}(s) \forall s \in S \). By Axiom 5, there is a transitive chain of equally preferable lotteries from \( \tilde{x} \) to \( \tilde{y} \). That is, let \( \tilde{x}' \) be a lottery which is otherwise as \( \tilde{x} \) except that the outcome of the lottery in state \( s \) has been replaced for the outcome of lottery \( \tilde{y} \). By Axiom 5 \( \tilde{x} \) and \( \tilde{x}' \) are equally preferable. Let us then change the outcome of \( \tilde{x}' \) in state \( s' \) to the outcome that \( \tilde{y} \) has in this state and denote the resulting lottery as \( \tilde{x}'' \). By Axiom 5 \( \tilde{x}' \sim_c \tilde{x}'' \). Let us repeat the procedure until all of the states in \( S \) have been gone through and we end up in lottery \( \tilde{y} \). By transitivity, \( \tilde{x} \sim_L \tilde{x}' \sim_L \tilde{x}'' \sim_L \cdots \sim_L \tilde{y} \). Q.E.D.

Note in particular that Proposition 2 can alternatively be formulated with the help of the DM’s utility function \( u \) as follows: \( u(\tilde{x}(s)) = u(\tilde{y}(s)) \forall s \in S \implies \tilde{x} \sim_L \tilde{y} \). It now becomes useful to introduce the concept of a utility lottery. A utility lottery describes the distribution of utility that can be acquired from a lottery. A utility lottery is formed by replacing each outcome of the lottery with the utility of the outcome. Figure 1 gives two examples of a lottery and its utility lottery. Mathematically, a utility lottery of a lottery \( \tilde{x} \) is a composition of utility function \( u \) with \( \tilde{x} \), i.e. the function \( u \circ \tilde{x} \), or when written using the notation that we commonly employ with random variables, \( u(\tilde{x}) \). Given that the equality sign between utility lotteries is defined as state-wise equality of utilities, Proposition 2 can be rewritten as follows using utility lotteries:

**PROPOSITION 2 (REWRITTEN):** \( u(\tilde{x}) = u(\tilde{y}) \implies \tilde{x} \sim_L \tilde{y} \).
Since two lotteries with the same utility lottery must be equally preferable and each lottery has a respective utility lottery, it is possible to define a preference relation on utility lotteries that is linked to the DM’s preferences over lotteries through an equivalence relation. Mathematically, this can be understood as partitioning $L$ into indifference classes, each represented by a single utility lottery, and then comparing these indifference classes (utility lotteries) to each other. This is useful for many reasons, in particular, because utility lotteries are real-valued as opposed to lotteries that may have non-numeric consequences. Indeed, sometimes it may be easier to assess the DM’s preferences over utility lotteries and then use this information to infer what his/her preferences over lotteries must be. This preference relation also turns out to play a major role in the separation result of this paper.

3.5 Indifference Classes

Figure 1: Two lotteries and their utility lotteries.

\[
u(x) = \sqrt{x}
\]
Indifference class relationship between a preference relation and the one on the set of implied indifference classes is defined as follows. Let $\succeq$ be a preference relation on a set $A$ and the implied set of indifference classes on $A$ be $\Pi$. The indifference class relationship between the preference relation on the set of indifference classes $\succeq_i$ and $\succeq$ on $A$ is then defined as follows:

$$a \succeq b \Leftrightarrow I(a) \succeq_i I(b) \quad \forall a, b \in A,$$

where $I(x) \in \Pi$ denotes the indifference class of $x$ in $A$. This simply implies that, if $a$ is preferred to $b$, then also the respective indifference class of $a$ is preferred to the indifference class of $b$.

For example, suppose that the DM has the following preferences over the following lotteries:

![Lotteries](image)

Obviously, the above is equivalent to the following comparison of indifference classes:

![Indifference classes](image)
Finally, since each of these indifference classes is represented by a single utility lottery, we can define the following preferences over utility lotteries, where the preference relation is to be understood as an *indifference class relationship* to preferences over lotteries:

![Diagram of indifference classes](image)

Following the above logic, we define the set of utility lotteries $Q = \{ u(\tilde{x}) \mid \tilde{x} \in L \}$, which by Proposition 2 represents elementary indifference classes on $L$. In order to distinguish the utility function with which the set $Q$ has been produced, it may be sometimes necessary to denote the set as $Q''$ instead. However, for the time being, we will drop the index.

**DEFINITION 5: Set of utility lotteries.** $Q = \{ u(\tilde{x}) \mid \tilde{x} \in L \}$

Let us then define a weak order $\succeq$ on $Q$.

**AXIOM 6: Weak ordering of utility lotteries.** *The DM’s preferences over $Q$ form a weak order* $\preceq$ on $Q$.

As discussed above, because $Q$ represents the set of elementary indifference classes on $L$, the relationship between $\preceq$ and $\succeq$ is defined as

**AXIOM 7: Indifference class relationship.** For all $\tilde{x}$ and $\tilde{y}$ in $L$, $\tilde{x} \succeq \tilde{y} \Leftrightarrow u(\tilde{x}) \preceq_Q u(\tilde{y})$.  

16
3.6 **Absolute Dominance**

*Consistency with strict absolute dominance* forms the final part of the axiomatic system. The axiom is based on the observation that there is a necessary consistency relation between the DM’s preferences over consequences and his/her preferences over lotteries: if the outcome of a lottery is better than the outcome of another lottery in each state of nature, the lottery should be preferred to the other. Putting this in more precise terms, it is required that

\[
\text{if } \bar{x}(s) \succeq_c \bar{y}(s) \quad \forall s \in S \quad \text{and the relation is strict at least for one } s, \quad \text{then } \bar{x} \succ_L \bar{y}.
\]

Note that, because \( u \) is the representation of \( \succeq_c \), \( \bar{x}(s) \succeq_c \bar{y}(s) \quad \forall s \in S \) can equivalently be expressed in a more practical form of \( u(\bar{x}(s)) \geq u(\bar{y}(s)) \quad \forall s \in S \).

For example, suppose that the DM prefers an apple to a carrot. Then, it is reasonable to assume that, for the following lotteries, the DM would prefer the first lottery:

The above consistency requirement can be conveniently expressed with the help of the absolute dominance (AD) relation. Given two real-valued acts \( X \) and \( Y \), the weak relation of AD is defined as

\[
X \succeq_{AD} Y \iff X(s) \geq Y(s) \quad \forall s \in S.
\]

Because by Axiom 7 preferences over utility lotteries are equivalent to preferences over
lotteries, we can replace \( \tilde{x} \succ_L \tilde{y} \) in the above consistency requirement with 
\( u(\tilde{x}) \succ Q u(\tilde{y}) \). Now, we can use the AD relation to define the consistency requirement as follows:

**AXIOM 8: Consistency with strict absolute dominance.** For all \( \tilde{x} \) and \( \tilde{y} \) in \( L \),
\[
 u(\tilde{x}) \succ_{AD} u(\tilde{y}) \Rightarrow u(\tilde{x}) \succ Q u(\tilde{y}).
\]

Continuing our earlier example, let us suppose that \( u(\text{apple}) = 20 \) and \( u(\text{carrot}) = 15 \). Then, Axiom 8 can be understood as the comparison of the following two utility lotteries where the first is “stochastically greater” than the second:

![Lottery Diagram](image)

Now, note that, using Axiom 7 and the absolute dominance relation, Proposition 2 can be rewritten as 
\[
 u(\tilde{x}) \sim_{AD} u(\tilde{y}) \Rightarrow u(\tilde{x}) \sim Q u(\tilde{y}).
\]
This observation together with Axiom 8 implies that \( \succsim_Q \) is fully consistent with \( \succsim_{AD} \), i.e. the relation is a *completion* of the absolute dominance relation. Thus, when coupled with Axiom 7, the present theory’s two consistency requirements between the outcomes of the lotteries (preferences for which are described by \( \succsim_C \)) and the DM’s preferences over lotteries (\( \succsim_L \)) imply a central consistency property for \( \succsim_Q \). Recognition of this property becomes useful in analyzing feasible representations for \( \succsim_Q \).

### 3.7 Separable Preference Model

We can now develop the preference model implied by Axioms 1–8. Let \( \tilde{x} \) be a lottery
and \( \tilde{\delta} \) be a degenerate lottery in \( L \) such that \( \tilde{\delta} \sim_L \tilde{x} \). By Axiom 4 (Existence of Equally Preferred Degenerate Lottery) we know that \( \tilde{\delta} \) does exist. For the sake of clarity, let us also denote \( X \equiv u(\tilde{x}) \) and \( \delta_{\tilde{x}} \equiv u(\tilde{\delta}) \). By Axiom 7 (Indifference Class Relationship), \( X \sim_Q \delta_X \). Note that \( \delta_X \) is unique, even though \( \tilde{\delta} \) is not necessarily so, because Axiom 8 (Consistency with Strict Absolute Dominance) implies that the DM cannot be indifferent between two different degenerate utility lotteries.

Importantly, we denote the single outcome of degenerate utility lottery \( \delta_X \) by \( CE[X] \), which is called the (utility) certainty equivalent of \( X \). \( CE \) can be interpreted as a functional from the set of utility lotteries to real numbers, called certainty equivalent operator, \( CE : Q \rightarrow \mathbb{R} \). For each utility lottery \( X \), the operator returns the amount of utility that the DM finds equally preferable to that utility lottery.

**Definition 6: Certainty equivalent operator.** The functional \( CE \) is called certainty equivalent operator and defined as \( CE : Q \rightarrow \mathbb{R} \) such that, for each \( X \) in \( Q \), \( CE[X] = c \), where \( c \) is the single outcome of a degenerate utility lottery \( \delta_X \) that is equally preferable to \( X \), i.e. \( X \sim_Q \delta_X \) and \( \delta_X(s) = c \forall s \in S \).

Note that from the transitivity of \( \succsim_Q \), it follows that \( X \succsim_Q Y \iff \delta_X \succsim_Q \delta_Y \), where \( Y \equiv u(\tilde{y}) \), \( \delta_Y \equiv u(\tilde{\delta})' \), and \( \tilde{\delta}' \sim_L \tilde{y} \). Furthermore, consistency with absolute dominance implied by Axiom 8 and Proposition 2 requires that \( \delta_X \succsim_Q \delta_Y \iff CE[X] \geq CE[Y] \). By combining these two observations, we get \( X \succsim_Q Y \iff CE[X] \geq CE[Y] \). We can then obtain the preference model for lotteries by using Axiom 7 (Indifference Class Relationship) and by substituting \( X \) and \( Y \) with their definitions above.

**Proposition 3: Separable preference model.** For all \( \tilde{x} \) and \( \tilde{y} \) in \( L \), \( \tilde{x} \succeq_L \tilde{y} \iff CE[u(\tilde{x})] \geq CE[u(\tilde{y})] \).

**Proof:** Proven as above.
Note in particular that since \( X \sim_\varnothing Y \iff CE[X] \geq CE[Y] \), \( CE[.] \) is a representation of \( \sim_\varnothing \). The general form of the DM’s preference functional over lotteries is \( U[\tilde{x}] = CE[u(\tilde{x})] \), which is also the representation of \( \sim_L \). A central aspect of this functional is that it is separated into two components that represent the two other preference relations of the theory: (i) the utility function \( u \), which is the representation of \( \sim_C \) and (ii) the certainty equivalent operator \( CE \), which represents \( \sim_\varnothing \). Since the DM’s utility function represents solely the DM’s preferences over sure outcomes, it logically follows that the CE operator must represent everything else about the DM’s preferences over lotteries, including his/her risk preferences.

Although we are now relatively close to the final goal of the paper, its accomplishment requires two further steps to be taken: Before the separation of the concepts of risk aversion and diminishing marginal utility can be established, we need to (i) define what is meant by risk aversion and (ii) introduce classic cardinal utility functions (Jevons (1871/1970)), as without cardinality there will be no such concept of diminishing marginal utility. We turn to this final part of the theory in Section 5 after examining certainty equivalent operators in more detail.

As a final remark, it is worth noting that the preference model does allow the use of state-dependent utilities, although at first glance the construction may seem to rule state-dependence out. Additional discussion on the topic is given in Appendix B.

### 3.8 Dependence of Certainty Equivalent Operator on Utility Function

A central observation regarding certainty equivalent operators is that the exact form of a certainty equivalent operator depends on the selected utility function. This is shown by Proposition 4. This is a particularly important observation, because utility functions are, in general, ordinal, so that if \( u \) is an admissible utility function, then also \( v(x) = \phi(u(x)) \), where \( \phi \) is a strictly increasing function, is as good a utility function as \( u \).
Nevertheless, once the CE operator is known under one utility function, it can be derived under all other utility functions, so that this property does not impose practical difficulties. In the following, we denote the certainty equivalent operator under utility function \( u \) by \( CE^u \).

**PROPOSITION 4: Dependence of certainty equivalent operator on utility function.** Let \( \phi \) be a strictly increasing function and \( u \) an ordinal utility function. Then, \( v(.) = \phi(u(.) ) \) is also an ordinal utility function and for all \( \bar{x} \) in \( L \) \( CE^u [v(\bar{x})] = \phi \left( CE^u \left[ \phi^{-1} (v(\bar{x})) \right] \right) \), i.e.

\[
CE^u [X] = \phi \left( CE^u \left[ \phi^{-1} (X) \right] \right).
\]

**PROOF:** Suppose that there is a degenerate lottery \( \tilde{\delta} \) such that \( \tilde{\delta} \sim \bar{x} \). The variable \( \tilde{\delta} \) has a single possible outcome \( \tilde{d} \) for which \( u(\tilde{d}) = CE^u \left[ u(\tilde{\delta}) \right] \). We can also see that

\[
CE^u [v(\tilde{\delta})] = v(\tilde{d}) = \phi(u(\tilde{d})) = \phi \left( CE^u \left[ u(\tilde{\delta}) \right] \right) .
\]

By equal preference we have

\[
CE^u [u(\bar{x})] = CE^u \left[ u(\tilde{\delta}) \right] \quad \text{and} \quad CE^u [v(\bar{x})] = CE^u \left[ v(\tilde{\delta}) \right] ,
\]

wherefrom we get

\[
CE^u [v(\bar{x})] = \phi \left( CE^u \left[ u(\bar{x}) \right] \right) = \phi \left( CE^u \left[ \phi^{-1} (v(\bar{x})) \right] \right) . \quad \text{Q.E.D.}
\]

### 4 SOME CERTAINTY EQUIVALENT OPERATORS

In principle, unless additional properties are necessary for a certainty equivalent operator in the application context, a certainty equivalent operator can be any functional that satisfies the Axioms 1–8. In particular, (i) for any degenerate utility lottery, the operator returns its single outcome (in choice under risk, this is equal to expectation), and (ii) it is consistent with absolute dominance (or with first-degree stochastic dominance, which implies consistency with AD). In Section 4.2 we show that this is the case for a mean-risk model used in Gustafsson and Salo (2005). Other potential candidates for a certainty equivalent operator include the expectation operator, transformations of the expectation operator of the form \( CE[X] = \phi^{-1} \left( E[\phi(X)] \right) \), and operators using a non-additive
capacity measure (Choquet-integrals). Perhaps the most important example of a certainty equivalent operator is given in the next section, where we examine how SEUT can be combined with triple preference theory.

4.1 SUBJECTIVE EXPECTED UTILITY THEORY

It is instructive to examine the certainty equivalent operator under expected utility theory. Therefore, let us assume next that, in addition to the axioms of triple preference theory, the DM abides by the axioms of subjective expected utility theory (Savage (1954)). We use specifically Savage’s (1954) framework, because it takes the same basic modeling approach as the present theory by using acts to model lotteries under uncertainty.

In stating the Savage axioms within the present framework, we use the following notation. Let $S$ be the set of events (sets of states in $S$) on $S$, forming a $\sigma$-algebra on $S$. The notation $\tilde{x}_E \tilde{y}$ is used to refer to a lottery in $L$ for which $\tilde{x}_E \tilde{y}(s) = \tilde{x}(s) \forall s \in E$ and $\tilde{x}_E \tilde{y}(s) = \tilde{y}(s) \forall s \notin E$, where $E$ is an event in $S$. An event $E$ is said to be null if for all for all lotteries $\tilde{x}$, $\tilde{y}$ and $\tilde{z}$, $\tilde{x}_E \tilde{z} \sim \tilde{y}_E \tilde{z}$. Axiom 3 corresponds to Savage’s axiom P1, so we start here at Savage’s axiom P2.

**AXIOM SEU-P2: Sure-thing principle.** For all $\tilde{x}$, $\tilde{x}'$, $\tilde{y}$ and $\tilde{y}'$ in $L$ and all $E$ in $S$, $\tilde{x}_E \tilde{y} \succ_L \tilde{x}_E \tilde{y}'$ if and only if $\tilde{x}_E \tilde{y} \succ_L \tilde{x}_E \tilde{y}'$.

**AXIOM SEU-P3: Eventwise monotonicity.** For all $\tilde{x}$ in $L$, all degenerate $\tilde{\delta}$ and $\tilde{\delta}'$ in $L$, and all non-null $E$ in $S$, $\tilde{\delta}_E \succ_L \tilde{\delta}'$ if and only if $\tilde{x}_E \succ_L \tilde{\delta}'$.

**AXIOM SEU-P4: Weak comparative probability.** For all $E$ and $E'$ in $S$ and all degenerate $\tilde{\delta}$, $\tilde{\delta}'$, $\tilde{\delta}''$ and $\tilde{\delta}'''$ in $L$ such that $\tilde{\delta}_E \succ_L \tilde{\delta}'$ and $\tilde{\delta}''_E \succ_L \tilde{\delta}'''$, $\tilde{\delta}_E \tilde{\delta}' \succ_L \tilde{\delta}_E \tilde{\delta}'$ if and only if $\tilde{\delta}'_E \tilde{\delta}'' >_L \tilde{\delta}''_E \tilde{\delta}'''$. 

22
AXIOM SEU-P5: Nondegeneracy. For some degenerate \( \tilde{\delta} \) and \( \tilde{\delta}' \) in \( L \), \( \tilde{\delta} \succ_L \tilde{\delta}' \).

AXIOM SEU-P6: Small event continuity. For all \( \tilde{x}, \tilde{y} \) and \( \tilde{z} \) in \( L \) such that \( \tilde{x} \succ_L \tilde{y} \), there is a finite partition \( (E_i)_{i=1}^n \) of \( S \) such that for all \( i \), \( \tilde{x} \succ_L \tilde{x}_i \) and \( \tilde{z}_i \succ_L \tilde{y} \).

When the above five axioms and Axioms 1-8 hold, we obtain the following subjective expected utility (SEU) representation theorem for the DM’s preferences over lotteries:

PROPOSITION 5: Subjective expected utility model. There exist some probability measure \( \mathbb{P} \) and some utility function \( u \) such that for all \( x \) and \( y \) in \( L \),

\[
\mathbb{P} \text{[} u(x) \text{]} \geq \mathbb{P} \text{[} u(y) \text{]},
\]

where \( \mathbb{P} \text{[} u(x) \text{]} = \int_S u(\tilde{x}(s))d\mathbb{P}(s) \). Further, \( u \) is cardinal in the sense that it is unique up to a positive affine transformation.


Taking the utility function from the SEU model in Proposition 5 and using it in Proposition 3, we obtain \( CE^u[X] = \mathbb{P} \text{[} X \text{]} \). Now, using Proposition 4, it is possible to derive a CE operator for a von Neumann-Morgenstern expected utility maximizer for utility functions that are not von Neumann-Morgenstern utility functions. Since \( u \) is a cardinal von Neumann-Morgenstern utility function, we will henceforth denote it by \( u_{vNM} \) to make a clear distinction to ordinal utility functions. Also, we drop the probability measure from the expectation operator’s subscript, as so far \( \mathbb{P} \) is the only probability measure that has been introduced. Indeed, let \( u_{vNM}(x) = \phi(\nu(x)) \), where \( u_{vNM} \) is a von Neumann-Morgenstern utility function, \( \phi \) is a strictly increasing function, and \( \nu \) is some ordinal utility function. Since by Proposition 5 we have \( CE^{u_{vNM}}[u_{vNM}(\tilde{x})] = \mathbb{P}[u_{vNM}(\tilde{x})] \), we get from Proposition 4

\[
CE^\nu[\nu(\tilde{x})] = \phi^{-1}\left( \mathbb{P}\left[ \phi(\nu(\tilde{x})) \right] \right).
\]

One may observe that the above result bears an interesting similarity to the formula of
certainty equivalent in expected utility theory. This result will become useful later in Section 5 where we examine risk attitudes.

4.2 MEAN - LOWER SEMI-ABSOLUTE DEVIATION MODEL

As mentioned previously, a combination of a mean-lower semi-absolute deviation (mean-LSAD) model and a multi-attribute value function is employed as a possible objective function in Gustafsson and Salo (2005). In this section, we show that this objective function is a valid representation of the DM’s preferences over lotteries under the present theory. This is accomplished by proving that the mean-LSAD model is consistent with first-degree stochastic dominance (FSD) when the risk aversion parameter is limited to a certain interval, for consistency with FSD implies consistency with AD. The reader can easily verify that the model equals expectation for all degenerate random variables. The weak relation of FSD is defined as \( X \succeq_{FSD} Y \iff F_X(\xi) \leq F_Y(\xi) \forall \xi \in \mathbb{R} \), where \( F_X(x) = \mathbb{P}\{X \leq x\} \) is the cumulative distribution function of random variable \( X \). LSAD (see e.g. Ogryczak and Ruszczynski (1999, 2001)) is expressed as

\[
LSAD[X] = \int_{-\infty}^{\mu_Y} (\mu_X - x) dF_X(x),
\]

where \( \mu_X \) denotes the expectation of \( X \).

As a byproduct, it is also shown that the mean-LSAD model is consistent with the second-degree stochastic dominance (SSD; see, e.g., Ogryczak and Ruszczynski 1999, 2001) by a pair of implications similar to FSD when the risk aversion parameter is limited to a smaller interval than in consistency with FSD. The weak relation of SSD is given by the equivalence

\[
X \succeq_{SSD} Y \iff \int_{-\infty}^{x} F_X(y) dy \geq \int_{-\infty}^{x} F_Y(y) dy \forall x \in \mathbb{R}.
\]

Consistency with SSD is sometimes an interesting property, because it implies that the DM is consistent with FSD and that he/she is either risk neutral or risk averse. Ogryczak and Ruszczynski (1999, 2001) have shown that the mean-LSAD model is consistent with SSD using a different approach; this latter part of the proposition can be regarded as an alternative (and perhaps simpler) derivation of their result.
**PROPOSITION 6**: Example using mean-lower semi-absolute deviation model. Let \( CE[X] = E[X] - \lambda \cdot LSAD[X] \). If \( \lambda \in [-1, 1] \), then \( CE \) is consistent with first-degree stochastic dominance, i.e., \( X \succ_{FSD} Y \Rightarrow CE[X] > CE[Y] \) and \( X \sim_{FSD} Y \Rightarrow CE[X] = CE[Y] \). Moreover, when \( \lambda \in [0, 1] \), \( CE \) is consistent with second-degree stochastic dominance.

**PROOF**: See Appendix C.

Results similar to Proposition 6 can be derived for mean-risk models with other risk measures, too. For example, a mean-risk model using absolute deviation (Konno and Yamazaki (1991)) is consistent with FSD when \( \lambda \in [-\frac{1}{2}, \frac{1}{2}] \) and with SSD when \( \lambda \in [0, \frac{1}{2}] \), because absolute deviation always equals twice the respective lower semi-absolute deviation. Ogryczak and Ruszczynski (1999, 2001) have also achieved similar results for general classes of mean-risk models, but these concern only consistency with SSD and higher degrees of stochastic dominance and are thus applicable for risk-averse decision makers only.

## 5 RISK ATTITUDES

We can now turn to the final part of the paper to establish the separation result: The concept of risk aversion itself. Although it would also be necessary to introduce cardinal utility in order to separate diminishing marginal utility from risk aversion, it happens to turn out that one needs to introduce cardinal utility already to obtain a reasonable definition of risk aversion.

### 5.1 LIMITATIONS OF CONVENTIONAL DEFINITIONS

Conventionally, risk aversion has been defined so that the DM is risk averse if, for each lottery, he/she is indifferent between the lottery and a sure outcome that is less than the expected value of the lottery (see e.g. French (1986) and Keeney and Raiffa (1976)). However, such definition implicitly assumes that (a) the outcomes are numeric, as
otherwise expectation of outcomes is not defined, and that (b) higher outcomes are preferred to lower outcomes, as otherwise a sure outcome that is lower than expectation is not necessarily worse than the expectation, which is necessary so that the definition implies that the DM is risk averse in the sense that he/she is willing to sacrifice something on expectation to avoid risk.

For example, it would be rather difficult to use the above definition in conjunction with the apples-and-oranges lotteries introduced in Section 3 even when the probabilities of the lotteries’ outcomes were known. Also, it would be foolish to say that a DM who is indifferent between a certain cost of $8 and a lottery of equally likely costs of $0 and $20 would be risk averse, as a cost of $8 is a better outcome than a cost of $10.

In recognition of the second shortcoming, Keeney and Raiffa (1976) have suggested that risk attitudes should be defined differently depending on what the DM’s preference ordering of the consequences (i.e. utility function) is. For example, they suggest that in the definition of risk aversion the inequality sign should be reversed if the preference order of consequences is decreasing. However, Keeney and Raiffa (1976) do not provide proper guidelines to come up with the definitions when the DM’s preferences are non-monotonic. Even if some such guidelines could be developed, it sounds rather inconvenient to have an infinite number of definitions for all the infinite kinds of utility functions that the DM may have, which raises the question whether the conventional definitions of risk attitudes are actually lacking something and whether it is possible to develop a single generic definition for risk aversion that would be independent of the DM’s preference ordering for consequences.

Because in multi-attribute decision making consequences are vectors, for which a one-dimensional expectation is not defined, Keeney and Raiffa (1976) have also tackled the first shortcoming by suggesting that the DM’s risk attitude should be defined separately on each attribute, giving rise to the idea of attribute-specific risk attitudes. This concept, however, is problematic in the general context: Attributes (dimensions of a vector / tuple) themselves may be non-numeric, in which case the attribute-specific risk attitude would
be undefined. Also, unless the DM’s utility function takes a specific decomposition form, there may not be attribute-specific utility functions that would be needed to calculate the DM’s certainty equivalents on the attributes.

5.2 Utility Scales

The two implicit assumptions underlying conventional definitions of risk attitudes are closely related to the concept of a *utility scale* (Krantz et al. (1971)). A utility scale is a numeric set where numbers represent the DM’s preferences; in such a set a higher number indicates that the associated consequence is preferred to a consequence with a lower utility number. Utility scales have exactly two key properties: they are numeric and higher numbers are better than lower numbers. Combining this observation with the two implicit assumptions, it can be concluded that, to make full conceptual sense, conventional definitions of risk attitudes require the underlying scale to be a utility scale. Note that any consequence scale which satisfies these two properties, such as money, is also a utility scale, generated by an admissible utility function \( u(x) = x \).

One can also invoke here the argument derived from the decomposition result in Proposition 3: Since the utility function represents the DM’s preferences under certainty and the CE operator represents everything about the DM’s preferences over lotteries that the utility function does not represent, it must be the CE operator that captures the DM’s risk attitude. It also happens that the CE operator is defined on the set of utility lotteries whose outcome space is, by definition, a utility scale, and therefore the outcomes of the CE operator are also utilities. This means that the DM’s risk attitude necessarily needs to be defined by using utility values (values of the CE operator) and utility-valued functions (utility lotteries).

5.3 Classic Cardinal Utility Functions

While it is fairly obvious that definition of risk attitudes implicitly requires the use of a utility scale, a more complex topic is the necessity of *cardinal* utility in the definition of risk attitudes. In the following, we define what we mean by cardinality and which kind of an axiomatic structure underlies classic cardinal utility functions.
A utility function \( u \) is said to be *cardinal* if for each other admissible utility function \( v \) there exist real numbers \( a > 0 \) and \( b \) such that \( v(x) = a \cdot u(x) + b \). Thus, a cardinal utility function can only be subjected to positive affine (linear) transformations, as opposed to ordinal utility functions that can be transformed using any strictly increasing function.

The classic cardinal utility function results from the measurement of the DM’s preferences over differences of consequences (Krantz et al. (1971); see also Jevons (1871/1970)), which is synonymous to the measurement of the intensity or strength of the DM’s preferences over consequences. Such preferences are defined using a preference relation specified on a set that is formed of pairs of consequences, \( C \times C \). We denote this preference relation by \( \succsim_C^\Delta \) and the pairs in \( C \times C \) by \( (x \leftarrow y) \), indicating a change from \( y \) to \( x \). Preference relation \( \succsim_C^\Delta \) can be defined using an axiomatic system called *algebraic difference structure*, which is specified in Appendix A (adapted from Krantz et al. (1971)). The axioms of this structure imply that there exists a real-valued function \( u \) such that for all \( c, d, e, f \) in \( C \), a change from \( d \) to \( c \) is preferred to a change from \( f \) to \( e \), denoted as \( (c \leftarrow d) \succsim_C^\Delta (e \leftarrow f) \) if and only if \( u(c) - u(d) \geq u(e) - u(f) \). That is,

\[
(c \leftarrow d) \succsim_C^\Delta (e \leftarrow f) \iff u(c) - u(d) \geq u(e) - u(f) \quad \forall c, d, e, f \in C
\]

Such \( u \) is a representation of the relation \( \succsim_C^\Delta \). Krantz et al. (1971) also show that such \( u \) is unique up to a positive affine transformation; i.e., such \( u \) is cardinal in the sense specified above. The relationship between \( \succsim_C^\Delta \) and \( \succsim_C \) is defined through the equivalence \( c \succsim_C d \iff (c \leftarrow d) \succsim_C^\Delta (d \leftarrow d) \), wherefrom it follows that \( c \succsim_C d \) if and only if \( u(c) \geq u(d) \). This implies that a utility function defined by \( \succsim_C^\Delta \) belongs to the set of admissible ordinal utility functions, but it is distinguished from other ordinal utility functions by its relationship to the relation \( \succsim_C^\Delta \), which makes it cardinal.
5.4 Need for Cardinal Utility

The central requirement for the use of a classic cardinal utility function comes from the preservation of the meaningfulness of the use of expectation in the definition of risk attitudes. Motivated by the observations in Sections 5.1 and 5.2 and the conventional definitions of risk attitudes (Keeney and Raiffa (1976)), let us suppose that risk neutrality is defined with respect to utility function \( u \) as follows, where expectation is taken with respect to some probability measure \( P \):

\[
\text{If } CE[X] = E[X] \text{ for all } X \text{ in } Q, \text{ then the DM is said to be risk neutral.}
\]

In particular, we can observe that under an ordinal utility function, the statement “\( CE[X] = E[X] \) for all \( X \) in \( Q \)” would not necessarily hold for all \( u \)’s if it holds for one. Indeed, if 
\[
CE^u[u(\bar{x})] = E[u(\bar{x})] \quad \text{under } u,
\]
then under \( v(\bar{x}) = \phi^{-1}(u(\bar{x})) \) we have, by Proposition 4,
\[
CE^v[v(\bar{x})] = \phi^{-1}\left(E[\phi(v(\bar{x}))]\right),
\]
so that \( CE \) will, by Jensen’s inequality, be equal to the expectation if and only if \( \phi \) is a (positive) affine transformation, i.e. the utility function is cardinal. In contrast, under ordinal utility the definition is dependent on the selected \( u \) and is thus meaningless. Therefore, in defining risk attitudes, it is necessary to limit the choice of utility function to those that are cardinal.

We can understand why this may be by observing a link between cardinal utility scales and the concept of dispersion, which can be regarded as synonymous with the concept of risk. For instance, we may observe that symmetric dispersion measures like variance assume that the underlying scale is such that equally large changes from expectation up- and downwards are equally preferable. Indeed, it does not make much sense to use variance to measure variability on a scale on which a minuscule change from the mean upwards is equally desirable to a very large change from below the mean to the mean. At a more conceptual level, we can conclude that since variance is based on measuring differences from the mean, these differences need to have a well-defined meaning for variance to have a well-defined meaning. In more general terms, we could say that meaningful measurement of dispersion / risk requires the specification of classic cardinal
utility, as without such utility, changes from one level to another will have no meaning and consequently there would be no meaningful measure of dispersion. And if there is no meaningful measure of risk, it is hardly surprising if there is no meaningful definition for risk aversion either.

Finally, it is good to bear in mind that if we are to separate the concepts of diminishing marginal utility and risk aversion, we need to introduce a classic cardinal utility function, as without cardinality there will be no such concept as diminishing marginal utility that could separated from risk aversion. However, only the above analysis makes the role of cardinal utility clear in the separation result.

5.5 Definitions

Taking the above observations together, we make the following definitions for risk attitudes, where expectation is taken with respect to some probability measure \( P \):

**Definition 7.** If the DM’s utility function is a representation of \( \succsim_c \) and \( CE[X] = E[X] \) for all \( X \) in \( Q \), then the DM is said to be risk neutral.

**Definition 8.** If the DM’s utility function is a representation of \( \succsim_c \) and \( CE[X] < E[X] \) for all nondegenerate \( X \) in \( Q \), then the DM is said to be risk averse.

**Definition 9.** If the DM’s utility function is a representation of \( \succsim_c \) and \( CE[X] > E[X] \) for all nondegenerate \( X \) in \( Q \), then the DM is said to be risk seeking.

Under these definitions, the DM’s risk attitude is preserved under change of the utility function, as proven by the following proposition. In the proof, the expression \( \pi^u[X] = E[X] - CE^u[X] \) denotes the DM’s (utility) risk premium under utility function \( u \).
**Proposition 7.** The DM’s risk attitude as defined by Definitions 7–9 does not depend on the selection of the utility function.

**Proof:** Since the DM’s utility function is a representation of $\succsim_C^\Delta$, it is cardinal, i.e. it can be subjected to positive affine transformations only. Let $f(x) = ax + b$, where $a > 0$ and $b$ are some constants. Let $v = f \circ u$ and $Y = f(X)$. By Proposition 4, we have $\pi'[Y] = E[Y] - CE'[Y] = E[Y] - f(CE^u[X]) = E[Y] - f(E[X] - \pi[X])$. By substituting $f$ with its definition, we get $\pi'[Y] = E[Y] - aE[X] - b + a\pi[X] = a\pi'[X]$. Thus, risk attitudes derived under $u$ are preserved under all $v$. Q.E.D.

In the special case where (i) consequences are numeric (such as money) and (ii) the utility function defined by $\succsim_C^\Delta$ is linear, Definitions 7–9 reduce to the conventional definitions of risk attitudes.

### 5.6 Implications

Definitions 7–9 constitute the final separation result of the paper: They formalize the role of a cardinal utility function and the certainty equivalent operator so that only the latter, which represents the part of the DM’s preferences that can reasonably be assumed to be related to risk aversion, determines the DM’s risk attitude. Even though the final separation result takes the form of three definitions, for derivation of these definitions we need the entire triple preference theory with the definition of set $Q$ of utility lotteries, the development of the CE operator, and the derivation of Propositions 3 and 4 that allow decomposition of the DM’s preference model into the CE operator and the utility function and define the dependence between these two components.

At the theoretical level, the main implication of Definitions 7–9 is that the DM’s risk attitude is related, in the spirit of Allais (1953, 1979a, 1979b), to the dispersion of classic cardinal utility, being irrespective of any attribute-specific utility functions that there might be. As to expected utility maximizers, recall that the von Neumann-Morgenstern utility function $u_{vNM}$ is a strictly increasing transformation of the classic cardinal utility function $u$ defined by $\succsim_C^\Delta$, i.e. $u_{vNM}(x) = \phi(u(x))$, where $\phi$ is a strictly increasing transformation.
function. Under $u$, a von Neumann-Morgenstern expected utility maximizer would therefore have the certainty equivalent operator $CE[X] = \phi^{-1}(E[\phi(X)])$. Thus, an expected utility maximizer’s risk attitude, when defined using Definitions 7–9, is related to the transformation function $\phi$ very much in the same way as one’s relative risk attitude is measured in Dyer and Sarin (1982).

At the practical level, the advantage of the present development is that the elicitation of the DM’s preference model can be conducted in two successive steps, each permitting a fair amount of complexity:

(i) elicitation of the DM’s utility function, and
(ii) elicitation of the DM’s certainty equivalent operator under that utility function.

This greatly simplifies the analysis of multi-attribute decision making under risk and uncertainty, because utility lotteries will always be real-valued and one-dimensional, and the analyst does not have to worry at the second stage anymore which kind of objects the consequences themselves are. Indeed, the consequences can, in general, be any mathematical entities, such as non-numeric $n$-tuples, so that the preference elicitation process even under certainty can be a challenging task. Also, there are several efficient techniques developed for the elicitation of value functions (see, e.g., French (1986) and Keeney and Raiffa (1976)) that can be now be used in the elicitation of the utility function. On the other hand, the certainty equivalent operator may be elicited by asking the DM to give certainty equivalents to a number of (utility) lotteries, and then finding the functional that best corresponds to the given certainty equivalents.

6 SUMMARY AND CONCLUSION

This paper developed a normative theory of choice under risk and uncertainty, called triple preference theory, that separates the concepts of risk aversion and diminishing marginal utility from each other. Underlying this separation result, there is a preference model for lotteries which is separated into the representation functions of the two other
preference relations of the theory. This decomposition holds under very mild conditions; in effect, rejection of either of the two central axioms (Axioms 5 and 8) would lead to rejection of expected utility theory as well. The decomposition form is formalized as follows:

\[ \text{For all lotteries } \tilde{x} \text{ and } \tilde{y} \text{ in set } L, \quad \tilde{x} \succeq_L \tilde{y} \iff CE[u(\tilde{x})] \geq CE[u(\tilde{y})], \]

where \( CE \) is the DM’s certainty equivalent operator and \( u \) is his/her a utility / value function. The CE operator is a functional that represents the DM’s preferences over “utility lotteries,” which are denoted by \( u(\tilde{x}) \) and \( u(\tilde{y}) \) and which belong to set \( Q \). The utility function, on the other hand, represents the DM’s preferences over sure outcomes, capturing the DM’s preferences under certainty. Hereby, it follows that it must be the CE operator that represents everything else about the DM’s preferences over lotteries, including the DM’s risk preferences.

The final separation result of the paper takes the form of a definition which is motivated by the observed limitations of conventional definitions of risk attitudes and the results of triple preference theory. In particular, we discuss that to make full conceptual sense, risk attitudes must be explored on a utility scale, and that in this context, meaningful definition with the expectation operator, as is conventional, necessitates the use of cardinal utility. Therefore, risk aversion is defined as follows, where preference relation \( \succeq_C^\lambda \) is represented by a classic non-stochastic cardinal utility function:

\[ \text{If the DM’s utility function is a representation of } \succeq_C^\lambda \text{ and } CE[X] < E[X] \text{ for all nondegenerate } X \text{ in } Q, \text{ then the DM is said to be risk averse.} \]

This means, in particular, that the DM’s risk attitude is related to the dispersion of classic cardinal utility, as called for by Allais (1953, 1979a, 1979b). As opposed to the conventional definition of risk aversion (Keeney and Raiffa (1976)), the above definition is, by construction, independent of the preference order of consequences and any possible
attribute-specific utility functions.

Importantly, the entire development is very generic, which makes it possible to use it together with several other theories of choice under risk and uncertainty: The axiomatic system does not assume the existence of a probability measure, nor does it assume any structure for consequences, wherefore it can immediately be employed in multi-attribute decision making under risk and uncertainty. Further, it allows for state-dependent utility and mean-risk-type preferences, and requires, in general, only the existence of an ordinal utility function. As an example, we showed that when the axioms of subjective expected utility theory are added on the top of those of triple preference theory, the DM’s risk attitude is captured by the transformation between the DM’s von Neumann-Morgenstern utility function and the DM’s classic cardinal utility function, which coincides with the postulate in Dyer and Sarin’s (1982) paper on relative risk aversion. A further example of the use of a preference functional allowed by triple preference theory can be found in Gustafsson and Salo (2005) where a combination of a mean-lower semi-absolute deviation model and a multi-attribute value function is used as a possible objective function in a multi-period project portfolio selection model.

From a practical viewpoint, the main implication of the paper is that, in choice under risk and uncertainty, the DM’s preferences can be elicited in two successive steps, first by eliciting the DM’s utility function, and then his/her certainty equivalent operator. This can make the elicitation of the DM’s preference model significantly more straightforward, as the certainty equivalent operator is defined on utility lotteries which are always one-dimensional and real-valued. Also, since the utility function can be elicited under certainty, it is possible to consider much more complex consequences and preference structures than what would be operationally possible if also risk considerations would be included in the elicitation process of the utility function.
7 REFERENCES


36


8 APPENDICES

8.1 APPENDIX A: ALGEBRAIC DIFFERENCE STRUCTURE

Suppose that the set $C$ is nonempty and $\succsim^\Delta_C$ is a weak order (complete and transitive) on $C \times C$. The pair $(C \times C, \succsim^\Delta_C)$ is an algebraic difference structure if and only if the following conditions are satisfied (Krantz et al. 1971):

1. Reversal: For all $a,b,c,d \in C$, $(a,b) \succsim^\Delta_C (c,d) \Rightarrow (d,c) \succsim^\Delta_C (b,a)$.

2. Monotonicity: For all $a,b,c,d,e,f \in C$,

   $\left((a,b) \succsim^\Delta_C (c,d) \land (b,e) \succsim^\Delta_C (d,f)\right) \Rightarrow (a,e) \succsim^\Delta_C (c,f)$.

3. Solvability: For all $a,b,c,d \in C$, if $(a,b) \succsim^\Delta_C (c,d) \succsim^\Delta_C (a,a)$, then there exist $e,f \in C$ such that $(a,e) \succsim^\Delta_C (c,d) \succsim^\Delta_C (f,b)$.

4. Archimedean property: If $a_1, a_2, \ldots, a_i \ldots$ is a strictly bounded standard sequence (i.e., $(a_{i+1}, a_i) \succsim^\Delta_C (a_2, a_1)$ for all $i$, and there exists $b,c \in C$ such that for all $i (c,b) \succsim^\Delta_C (a_i, a_i) \succsim^\Delta_C (b,c)$), then it is finite.

8.2 APPENDIX B: STATE-DEPENDENT UTILITY

In choice under risk, a lottery yields different outcomes, because a different state of nature realizes. However, apart from influencing the outcome of the lottery, a state of nature can affect the desirability of consequences to the DM. If so, we speak of state-dependent utility. This section shows briefly how state-dependent utilities can be implemented in the present theory.

The key is to consider the consequences in each state as individual consequences that are distinct from consequences of other states similarly as one does in the state-preference approach (see, e.g., Arrow (1965), Debreu (1959), and Hirschleifer (1965, 1966)). For example, “an apple in state 1” and “an apple in state 2” are two different consequences.
Suppose that \( S \) is a set of states of nature and that the set of consequences pertaining to state \( s \in S \) is \( C_s \). This set is a partition of \( C \) and it is disjoint from other similar partitions of \( C \). We require random consequences to assume only consequences in \( C_s \) in each state \( S \). This takes the form of an additional assumption:

**AXIOM A1: State-specific consequences.** \( \bar{x}(s) \in C_s \ \forall s \in S \ \forall \bar{x} \in L \).

Let \( x_s \in C_s \) denote the representative of consequence \( x \) (e.g., “an apple”) in state \( s \) (“apple in state \( s \)”). Then, the DM’s (state-dependent) utility function in state \( s \) can be defined as \( u_s(x) = u(x_s) \).

Under Axiom A1, the DM’s preferences over consequences, his/her utility function, and the rest of the theory can be developed from this standpoint exactly as before. Since the utility function can be defined over multi-attribute consequences, we can observe that utility can be state-dependent also when the decision problem is characterized by multiple attributes, which appears to contradict one of the conclusions in Miyamoto and Wakker (1996).

### 8.3 APPENDIX C: PROOF OF PROPOSITION 6

The equivalence implication of both FSD and SSD is trivial, since they hold if and only if \( F_X = F_Y \). Thus, let us examine the strict preference implication. Suppose first that \( X \) strictly dominates \( Y \) by FSD, i.e., that \( F_X(x) \leq F_Y(x) \ \forall x \in \mathbb{R} \) and that the inequality is strict at least for one \( x \).

Let us recall that the definition of infinite integrals \((a < c < b)\) is
\[
\int_a^b f(x) \, dx = \lim_{a \to -\infty} \int_a^c f(x) \, dx + \lim_{b \to +\infty} \int_c^b f(x) \, dx .
\]
Suppose that, for some \( a \) and \( b \), \( F_X(a) = F_Y(a) = 0 \) and \( F_X(b) = F_Y(b) = 1 \). This can be done without loss of generality, because we can eventually let, in the spirit of infinite integrals, \( a \) and \( b \) approach minus infinity and infinity, respectively. Then, by integrating
the expectation by parts we get

\[
E[X] = \mu_X = \int_a^b x dF_X(x) = \int_a^b x F_X(x) dx - \int_a^b F_X(x) dx = b - \int_a^b F_X(x) dx.
\]

The difference of expectations of \(X\) and \(Y\) is thus given by

\[
E[X] - E[Y] = \mu_X - \mu_Y = \int_a^b F_Y(x) dx - \int_a^b F_X(x) dx.
\]

Since \(F_X\) is smaller than \(F_Y\) by FSD, \(E[X]\) must be greater than \(E[Y]\). Also, the same holds if \(X\) dominates \(Y\) by the rules of SSD. This follows directly from setting \(x = b\) in Equation (2). Let us next examine the relationships between lower semi-absolute deviations of \(X\) and \(Y\). LSAD in Equation (1) can be written as follows:

\[
LSAD[X] = \int_a^{\mu_X} (\mu_X - x) dF_X(x) = \mu_X \int_a^{\mu_X} dF_X(x) - \int_a^{\mu_X} x dF_X(x) = \\
\mu_X F_X(\mu_X) - \left[ \int_a^{\mu_X} x F_X(x) dx - \int_a^{\mu_X} F_X(x) dx \right] = \int_a^{\mu_X} F_X(x) dx
\]

The difference of LSADs of \(X\) and \(Y\) is thus given by

\[
LSAD[X] - LSAD[Y] = \int_a^{\mu_X} F_X(x) dx - \int_a^{\mu_Y} F_Y(x) dx.
\]

The difference of certainty equivalents of \(X\) and \(Y\) is \(CE[X] - CE[Y] = E[X] - E[Y] - \lambda (LSAD[X] - LSAD[Y])\). By substituting the formula of LSAD difference we get

\[
CE[X] - CE[Y] = \mu_X - \mu_Y - \lambda \left[ \int_a^{\mu_X} F_X(x) dx - \int_a^{\mu_Y} F_Y(x) dx \right].
\]

Obviously if \(\lambda\) is zero, the consistency implication for both FSD and SSD holds, because certainty equivalent is reduced to expectation. Next, suppose \(\lambda\) is negative. Let us write \(m = -\lambda\) and both add and subtract \(m(\mu_X - \mu_Y)\) from the right-hand side of the equation. By combining the negative term with \(\mu_X - \mu_Y\), which results in \((1-m)(\mu_X - \mu_Y)\), and the positive term with the latter integral, we get

\[
CE[X] - CE[Y] = (1-m)(\mu_X - \mu_Y) + m \int_{\mu_Y}^b F_Y(x) dx - \int_{\mu_X}^b F_X(x) dx
\]

The \(\mu_X - \mu_Y\) component of the first term is positive, because difference of expectations in positive. The integral component of the second term is also positive if FSD holds. This results from that (i) FSD ensures \(F_X\) is smaller than \(F_Y\) for all \(x\) and that (ii) the integral over \(F_X\) has a shorter integration interval than that over \(F_Y\), since \(\mu_X > \mu_Y\). SSD, on the other hand, says nothing about the relative values of integrals whose lower bounds differ.
from minus infinity (i.e., here from $a$). The difference of certainty equivalents is thus positive when $m$ belongs to $[0,1]$, i.e., when $\lambda$ belongs to $[-1,0]$, and $X$ dominates $Y$ by FSD.

Next, suppose now that $\lambda$ is positive and make the following modifications to the formula of certainty equivalent difference:

$$CE[X] - CE[Y] = \mu_X - \mu_Y - \lambda \left[ \int_{\mu_X}^{\mu_Y} F_X(x)dx - \int_{\mu_Y}^{\mu_X} F_Y(x)dx \right] =$$

$$\mu_X - \mu_Y - \lambda \left[ \int_{\mu_X}^{\mu_Y} [F_X(x) - F_Y(x)]dx + \int_{\mu_Y}^{\mu_X} F_X(x)dx \right] =$$

$$\mu_X - \mu_Y + \lambda \left[ \int_{\mu_Y}^{\mu_X} [F_Y(x) - F_X(x)]dx - \int_{\mu_Y}^{\mu_X} F_X(x)dx \right]$$

The following inequalities clearly hold:

$$\mu_X - \mu_Y + \lambda \left[ \int_{\mu_Y}^{\mu_X} [F_Y(x) - F_X(x)]dx - \int_{\mu_Y}^{\mu_X} F_X(x)dx \right] \geq$$

$$\mu_X - \mu_Y + \lambda \left[ \int_{\mu_Y}^{\mu_X} [F_Y(x) - F_X(x)]dx - (\mu_X - \mu_Y)F_X(\mu_X) \right] \geq$$

$$\mu_X - \mu_Y + \lambda \left[ \int_{\mu_Y}^{\mu_X} [F_Y(x) - F_X(x)]dx - (\mu_X - \mu_Y) \right] =$$

$$(1 - \lambda)(\mu_X - \mu_Y) + \lambda \left[ \int_{\mu_Y}^{\mu_X} [F_Y(x) - F_X(x)]dx \right]$$

Again, $\mu_X - \mu_Y$ component of the first term is positive when FSD or SSD holds. However, the latter term now contains integrals with an infinite lower bound (i.e., $a$, which we let go to minus infinity), wherefore both FSD and SSD ensure that it is positive. Thus, both the terms above are positive and hence the difference of certainty equivalents is positive when $\lambda$ belongs to $[0,1]$ and $X$ dominates $Y$ by the rules of FSD or SSD.

To conclude, $CE$ is consistent with FSD when $\lambda \in [-1,1]$ and with SSD when $\lambda \in [0,1]$. Q.E.D.