Is Cumulative Prospect Theory a Serious Alternative for the Expected Utility Paradigm?

Lewandowski, Michal

European University Institute

14 November 2006

Is Cumulative Prospect Theory a Serious Alternative for the Expected Utility Paradigm?

Michal Lewandowski
EUI Florence, Italy

November 14, 2006

1 Introduction

The purpose of this paper is to demonstrate that Cumulative Prospect Theory is a serious alternative for Expected Utility Theory. It does not contradict Expected Utility, but includes it as a special example. A very **useful example**, because simple and yet very flexible, Expected Utility proved indispensable in many areas of economic analysis. Though a **special example**, because it does not capture some important effects observed in real choice behavior.

This paper is organized as follows.[1] In section 2, we try to outline the most important elements not embodied in Expected Utility Model, but systematically found in observed pattern of choices among decision makers. In section 3, we describe briefly the body of Prospect Theory, a very influential, early alternative to Expected Utility. Influential, because it is the first important contribution which tries to build a bridge between psychology and traditional economics. It initiated the whole wave of papers investigating psychological motives underlying decision processes. In section 4, we start by stating the problem with Prospect Theory as a way to rigorously model decision making - the difficulties in formalizing the editing phase of Prospect Theory and more importantly, the possibility of non-monotonicity. The Rank Dependent model is shown to solve the non-monotonicity problem of Prospect Theory. Furthermore, it is

[1] In this paper, we decided not to adopt theorem-proof writing style since we want to focus on intuition of results rather than mathematical rigor. Besides, most of the results presented in this article were already proved before.

argued that Rank Dependence model and the underlying assumptions are not just merely a technical tool, but there is a simple intuition behind. In section 5, we finally arrive at Cumulative Prospect Theory - the idea, which now in view of the models introduced in the preceding sections, makes perfect sense. The Cumulative Prospect Theory is shown to combine core elements of Prospect Theory, Rank-Dependent models, and additionally sign-dependence, which is a novel feature of the new model. Section 6 describes risk attitudes in Cumulative Prospect Theory. The aim here is to show, that the new theory is perfectly capable of incorporating any desired risk attitude, and more importantly, that these risk attitudes can be represented in a simple way. We describe stochastic dominance in Cumulative Prospect Theory, probability weighting issues and loss aversion. In section 7 we give some examples of economic phenomena explained by the new theory, try to point out some important environments in which we can expect the new model to produce better results than Expected Utility Theory and conclude.

2 Expected Utility Model - Critique

[2] The choice under uncertainty or risk is a fundamental issue in economics. Risk and uncertainty refer to situations where the decision-maker is faced with randomness. In case of risk there is some objective distribution over this randomness, whereas in case of uncertainty there is no. Without going into philosophical disputes, asking what exactly is this objective distribution, we simply assume that this is something given. It means we are given some reliable probabilities of events.[3]

In case of choice under risk, it was von Neumann and Morgenstern (1947) who demonstrated how a set of apparently reasonable axioms on preference can be shown to imply Expected Utility model (EU).

Savage (1954) demonstrated the same for the case of uncertainty providing grounds for Subjective Expected Utility model (SEU). The two models are basically equivalent, with one difference - in case of risk we are given the objective distribution and in case of uncertainty we build instead a distribution over subjective beliefs, which obey the properties of probability measure.

[2]This section draws on the following articles: Starmer (2000), Machina (2005), Nau (2004).

[3]We can think of rolling a dice, throwing a coin or playing a roulette game as situations in which we are given some objective distribution over the outcomes. On the other hand, the outcome of a football game or some military conflict might rather be regarded as uncertain, in which case we are not given any objective probability distribution over the outcome space.

Let us concentrate for the time being on the former case and more specifically we assume that that the object of choice is finite outcome lotteries of the form:

$$\mathbf{p} \equiv (x_1, p_1; ...; x_n, p_n) \tag{1}$$

assigning outcome x_i with probability p_i.

Ordering (transitivity and completeness) and continuity axioms together imply that the preferences over lotteries can be represented by a function $V(.)$, which assigns a real-valued index to each lottery. The crucial axiom of EU model is **independence**. It requires that for all lotteries $\mathbf{p}, \mathbf{q}, \mathbf{r}$: if $\mathbf{p} \succcurlyeq \mathbf{q}$ then $(\mathbf{p}, p; \mathbf{r}, 1-p) \succcurlyeq (\mathbf{q}, p; \mathbf{r}, 1-p)$. If independence holds together with ordering and continuity, then preferences over lotteries, such as in (1) can be represented by:

$$V_{EU}(\mathbf{p}) = \sum_{i=1}^{n} p_i u(x_i) \tag{2}$$

where $u(.)$ is the so called von Neumann Morgenstern utility function defined on the set of outcomes. The main strength of EU representation is its simplicity. Before Expected Utility was introduced, people tended to accept expected value representation. However, it leads to Saint Petersburg paradox and does not allow for variable risk attitudes. In EU both these problems are circumvented and more importantly, its mathematical properties allow very simple, straightforward, and yet very flexible representation of risk. Let me give you few examples:

- $V_{EU}(.)$ exhibits first order stochastic dominance preference iff $U(x)$ is an increasing function of x
- $V_{EU}(.)$ exhibits risk aversion iff $U(x)$ is a concave function of x
- $V^*_{EU}(.)$ is at least as risk averse as $V_{EU}(.)$ iff its utility function $u^*(.)$ is a concave transformation of $u(.)$

It is well known that the Expected Utility paradigm is very well established and there are many extremely important contributions based on this paradigm. To question this paradigm is therefore not an easy task. Economists have tried it already for fifty years, but Expected Utility still prevails. The main reason, that this theory is so popular is its parsimony. Very simple assumptions allow huge flexibility and produce results consistent with observed behavior in a wide variety of economic environments. All known alternatives for Expected Utility does not achieve this level of parsimony, thus to make a strong case for at least

one of them, we have to show that there are many environments in which an alternative theory works fine and Expected Utility doesn't work.
The Expected Utility, meaning both Expected Utility Theory and Subjective Utility Theory, critique which follows can be divided into three broad classes:

- Violations of independence axiom
- Violations of descriptive and procedural invariance
- Source dependence

2.1 Violations of independence axiom

Violations of the independence axiom, as documented by a huge experimental and empirical literature usually fall within two broad groups:

- Common consequence effect
- Common ratio effect

Common consequence effect
Suppose we have four compound lotteries:

$$\mathbf{b_1} \equiv (x, p; \mathbf{P}^{**}, 1-p) \quad \mathbf{b_2} \equiv (\mathbf{P}, p; \mathbf{P}^{**}, 1-p)$$
$$\mathbf{b_3} \equiv (x, p; \mathbf{P}^{*}, 1-p) \quad \mathbf{b_4} \equiv (\mathbf{P}, p; \mathbf{P}^{*}, 1-p)$$

where $\mathbf{P}$ involves positive outcomes both greater and less than x, and $\mathbf{P}^{**}$ first order stochastically dominates (FOSD) $\mathbf{P}^{*}$.
There is strong evidence in the literature that people often follow the pattern of choice, which reveals the following preferences: $\mathbf{b_1} \succ \mathbf{b_2}, \quad \mathbf{b_4} \succ \mathbf{b_3}$. However the independence axiom requires the following:

$$x \succcurlyeq \mathbf{P} \Rightarrow \mathbf{b_1} \succcurlyeq \mathbf{b_2},\ \mathbf{b_3} \succcurlyeq \mathbf{b_4}$$
$$x \preccurlyeq \mathbf{P} \Rightarrow \mathbf{b_1} \preccurlyeq \mathbf{b_2},\ \mathbf{b_3} \preccurlyeq \mathbf{b_4}$$

which is obviously in contradiction with the revealed preferences stated above.[4]
Common ratio effect
Suppose we have four lotteries:

$$\mathbf{c_1} \equiv (x, p; 0, 1-p) \qquad \mathbf{c_2} \equiv (y, q; 0, 1-q)$$
$$\mathbf{c_3} \equiv (x, \alpha p; 0, 1-\alpha p) \quad \mathbf{c_4} \equiv (y, \alpha q; 0, 1-\alpha q)$$

[4] The famous Allais paradox is a special case of common consequence effect.

where $p > q$ and $0 < x < y$.
There is strong evidence in the literature that people often follow the pattern of choice which reveals the following preferences: $\mathbf{c_1} \succ \mathbf{c_2}, \quad \mathbf{c_4} \succ \mathbf{c_3}$. Observe that $\mathbf{c_3}$ and $\mathbf{c_4}$ can both be written as compound lotteries in the following manner: $\mathbf{c_3} = (\mathbf{c_1}, \alpha; 0, 1-\alpha),\ \mathbf{c_4} = (\mathbf{c_2}, \alpha; 0, 1-\alpha)$.
Given that, the independence axiom requires:

$$\mathbf{c_1} \succcurlyeq \mathbf{c_2} \Rightarrow \mathbf{c_3} \succcurlyeq \mathbf{c_4}$$
$$\mathbf{c_1} \succcurlyeq \mathbf{c_2} \Rightarrow \mathbf{c_3} \succcurlyeq \mathbf{c_4}$$

which is again in contradiction with the revealed preferences stated above.
But for one difference, both common consequence and common ratio effect can be shown in the same manner as above for lotteries with negative outcomes. The difference is that revealed preference for lotteries with negative outcomes is everywhere reversed in comparison to lotteries with positive outcomes ($\preccurlyeq$ instead of $\succcurlyeq$). This phenomenon, that changing the sign of outcomes changes the revealed preference everywhere, is called the **reflection principle**.

2.2 Violations of descriptive and procedural invariance

Whereas economists usually agree that the theory of choice should account for violations of independence, they are not so unanimous about violations of descriptive and procedural invariance and violations of transitivity, monotonicity (first order stochastically dominating lotteries should be preferred to lotteries which they dominate) and completeness. Behavioral economists usually focus on descriptive theory and as such, they are willing to include any serious violation of classical theory assumptions as long as it allows them to describe real people behavior. "Orthodox" economists are more interested in normative aspects of the theory and accept certain axioms as long as they sound reasonable. Reasonable in a sense, that by violating it consciously, a decision maker would contradict the principle of maximizing the given objective, which he/she is assumed to maximize.
Violations of completeness, transitivity and monotonicity, although very likely to occur in the real world, didn't get much attention in economics literature, since it is believed that economics is amenable to mathematical description and without ordering axioms we cannot even define the meaningful concept of preference. Economists usually also agree, that violations of monotonicity could at most be a mistake in individual evaluation and not something which systematically happens in choice decisions.

That leaves us with descriptive and procedural invariance. It gained more attention in economics literature because of the seminal paper by Kahnemann and Tversky (1979). We can distinguish here several points:

- **framing effects** - offering a gain or a loss contingent on the joint occurrence of n independent events with probability p often gives different responses than offering the same loss or gain contingent on the occurrence of a single event with probability p^n
- **description form** - different responses depending on the description form of a lottery: matrix form, decision tree, roulette wheels, written statements, etc.
- **gamble or insure** - different responses for identical problems but framed either whether to gamble or whether to insure
- **response mode effects** - different responses depending on whether an experiment was designed to elicit certainty equivalent, gain equivalent or probability equivalent (which under EUT all should yield equivalently assessed utility functions)
- **preference reversal** - given two lotteries: the so called \$-bet $(X, p; 0, 1-p)$, and the so called P-bet $(x, P; 0, 1-P)$ where $X > x$ and $P > p$, people usually choose the P-bet but assign higher certainty equivalent to the \$-bet
- **reference dependence** - this is one of the two building blocks of Kahnemann and Tversky (1979) seminal contribution - it states that people usually do not assess final asset positions but they assess each outcome relative to some reference point and therefore they code outcomes as gains and losses

2.3 Source dependence

Source dependence concerns the distinction between risk and different kinds of uncertainty. Some economists argued that the expectation principle can be applied to decision under risk, where probabilities are known but not to decision under uncertainty or ignorance where probabilities are not known. There is strong evidence in the literature that agents' preferences depend not only on the degree of uncertainty but also on the source of uncertainty. This phenomenon, together with the problem of nonexistence of probabilistic beliefs, can be illustrated by Ellsberg paradox (Ellsberg, 1961). Suppose we have an urn with 30 red (R) balls and 60 other balls, either black (B) or yellow (Y). So there is 90 balls in the urn and the experiment is to choose one of them. Now consider four acts, where an act is the equivalent of a lottery in case of uncertainty - instead of probabilities of outcomes, we are given events, each of them yielding a particular outcome:

$$\begin{aligned} f_1 &\equiv (100,\mathrm{R};\quad 0,\mathrm{B};\quad 0,\mathrm{Y}), \\ f_2 &\equiv (\quad 0,\mathrm{R}; 100,\mathrm{B};\quad 0,\mathrm{Y}) \\ f_3 &\equiv (100,\mathrm{R};\quad 0,\mathrm{B}; 100,\mathrm{Y}) \\ f_4 &\equiv (\quad 0,\mathrm{R}; 100,\mathrm{B}; 100,\mathrm{Y}) \end{aligned}$$

It is commonly observed that people usually choose f_1 against f_2 and f_4 against f_3. However such preferences are inconsistent with any assignment of subjective probabilities $\mu(\mathrm{R}), \mu(\mathrm{B}), \mu(\mathrm{Y})$. To see this notice that if an individual were choosing according to SEU, then we could infer from the first choice that: $\mu(\mathrm{R}) > \mu(\mathrm{B})$ and from the second choice that: $\mu(\mathrm{R}\cup\mathrm{Y}) < \mu(\mathrm{B}\cup\mathrm{Y})$ and because probabilities sum to one: $1-\mu(\mathrm{B}) < 1-\mu(\mathrm{R})$. Hence $\mu(\mathrm{B}) > \mu(\mathrm{R})$, which contradicts the first choice. A preference for acts based on probabilistic partitions over acts based on subjective partitions is called **ambiguity aversion**. This is an important example of source dependence. People prefer to choose from the known distribution, rather than from the unknown one, although there is no objective reason why they should expect the unknown distribution to be less favorable.

3 Prospect Theory

[5] This section will briefly sketch the outline of Prospect Theory (PT) (Kahnemann and Tversky, 1979), as one of the two building blocks of Cumulative Prospect Theory.[6]

Kahnemann and Tversky (1979) started with stating the three basic tenets of Expected Utility Theory (EUT):

- the expectation principle: $V(x_1,p_1;...;x_n,p_n) = \sum_{i=1}^{n} p_i u(x_i)$
- asset integration: $(x_1,p_1;...;x_n,p_n)$ is acceptable at asset position w iff $V(w+x_1,p_1;...;w+x_n,p_n) > u(w)$

[5] This section is based on Kahnemann and Tversky (1979).

[6] Prospect Theory is only for decisions under risk. Cumulative Prospect Theory was extended to the case of uncertainty as well.

- risk aversion: $u(.)$ is a concave function of its argument

Then they presented evidence against these three basic tenets. Although this evidence is just a part of critique presented in the previous section, it is worth underscoring the main body of this evidence, because it gives direct motivation for the Prospect Theory. There are two basic effects emphasized in the paper:

- **Certainty effect** - can be shown both as a special case of common consequence and common ratio effect. It shows that people tend to violate the expectation principle if they are to choose between the lottery with certain outcome and the lottery with uncertain outcomes.
- The other effect occurs with lotteries in which it is **highly unlikely but possible to win**. In such circumstances, people tend to choose larger gain / smaller probability lottery, which often contradicts Expected Utility.

Both these effects indicate that people distort probability scale. The reason may be that the intuition of extreme probabilities like 0.001 or 0.999 is hard to grasp psychologically. Throwing a coin can give you quick intuition for 0.5 probability, but not for 0.001. The other reason may be that people intentionally pay more attention to extreme events and they choose as though they are implicitly assigning different-than-objective probabilities in their decision making. Or finally, it may be that since we usually make thousands of small decisions every day, we tend to use simple heuristics which simplify any given choice problem at hand. We may round some probabilities and neglect the others depending on our quick perception, and we may pay more attention to the possibility of big changes in our status quo, which in turn happen usually with small probability. Whatever the reason, the fact is that we tend to **overweight small probabilities of extreme events and underweight probabilities of moderate events.**

There are two other crucial points of Kahnemann and Tversky (1979) critique of EUT.

The first concerns the so called **probabilistic insurance**[7]. According to such an insurance scheme, you pay half of the regular premium. In case of damage, there is 50 per cent chance that you pay the other half of the premium and your losses are covered, and there is 50 per cent chance that you get back your insurance payment and suffer all the losses. The evidence suggests that people prefer regular insurance over probabilistic one. However, EUT predicts the opposite.

[7] It should not be confused with partial insurance.

The other systematic violation of the EU paradigm pointed out by Kahnemann and Tversky (1979) is the **isolation effect**. This effect occurs when the decision problem is sequential. In this case, people tend to ignore previous stages when making a decision at subsequent stages. This violates the EU basic supposition, that people evaluate final asset positions and the only probabilities that matter are the probabilities of final states. Isolation effect suggests that people evaluate outcomes of a lottery relative to some reference point, which usually correspond to the status quo. If the problem is sequential, the status quo of a decision maker changes after each stage and the subsequent stages are evaluated relative to a different reference point.
After the critique of Expected Utility Theory, Kahnemann and Tversky (1979) presented their model. According to it, the choice process involves two phases: editing and evaluation of prospects.

- **Editing** is meant to serve as a preliminary analysis of a prospect.[8] It specifies rules how to simplify a problem, it involves defining a reference point and hence deciding what is to be regarded as losses and what as gains, and possibly detecting dominance. This phase is needed to avoid some basic inconsistencies in choice.[9]
- **Evaluation** follows certain rules derived from observed agents' behavior. The most important two contributions are:
  - **reference dependence** - the carriers of value are gains and losses, which are perceived and hence evaluated differently, and not final assets
  - **decision weights** - nonlinear distortion of probability scale

There are two main conditions imposed on the shape of utility function and probability distortion in prospect theory, which follow from observed behavior:

- The first is **loss aversion**, according to which losses loom larger than gains. It means that utility function for losses is steeper than for gains,

[8]Prospect Theory and Cumulative Prospect Theory use the notion of prospect and value function instead of lottery/act and utility function, respectively, to underscore the reference dependence aspect of this theory. Prospects involve gains and losses instead of final assets.

[9]In Cumulative Prospect Theory, the authors abandoned the idea of editing phase. The reason for this is that it is difficult to formalize it, especially because the order of actions taken in this phase can have effects on what form of prospect survives until evaluation phase. However, as Kahnemann and Tversky (1979) emphasize, this phase plays important role in the decision making process and it can account for some oddities in observed choices.

and in particular there is a kink at the reference point. We will see in one of the subsequent sections that this kink plays important role in measuring loss aversion.

- The second is **diminishing sensitivity**, which implies that the impact of change in a given variable diminishes with a distance from the reference point. In terms of utility function, it means that it is concave for gains and convex for losses with $u(0) = 0$. In terms of probability distortion, there are two natural reference points -boundaries. One is certainty (probability one) and the other is impossibility (probability zero). The principle of diminishing sensitivity in this case means that the probability distortion (which is a continuous, strictly increasing function $w(.)$ on $[0, 1]$ domain with $w(0) = 0$ and $w(1) = 1$) has an inverse S-shape. The impact of a given change in probability diminishes with its distance from the boundary.

The implicit assumption of prospect theory is that the decision weight assigned to a specific outcome depends only on the probability of this outcome.[10] We will see in the next section that this assumption generates certain problems. It does not always satisfy stochastic dominance, which is regarded as highly undesirable by most economists.[11] And also, it is not readily extended to lotteries with a large number of outcomes. However the Rank Dependent model presented in the next section circumvents this problem. Kahnemann and Tversky (1992) in turn combine Prospect Theory and Rank Dependent model into one - Cumulative Prospect Theory, which will be introduced later in this paper.

4 Rank Dependence-Intuition

[12] The main problem with Prospect Theory is that it does not always satisfy stochastic dominance[13] and it is widely believed that any satisfactory theory should satisfy it. In this section we demonstrate that the central assumptions underlying the Rank Dependent models solve the problem posed by prospect theory and at the same time follow from intuitive arguments. We shall focus here on the case of uncertainty and not risk, since the analysis is parallel for

[10] Alternatively for the case of uncertainty only on the event generating this outcome.

[11] Machina (1983, p.97) argues that any theory, which fails to guarantee monotonicity is "in the author's view at last, unacceptable as a descriptive or analytical model of behavior."

[12] This section draws on Diecidue and Wakker (2001).

[13] See assumption A2 below.

both cases and uncertainty offers more generality.
Let S be an exhaustive mutually exclusive set of states of nature. Acts are finite-valued functions $f : S \rightarrow R$ where R is an outcome space. We denote $f \equiv (x_1, E_1; ...; x_n, E_n)$ where an event E_i yields outcome x_i. Events $E_i; i = 1, ..., n$ form a partition of the state space. We impose the following assumptions:

A1: (General Weighting Model) An act $(x_1, E_1; ...; x_n, E_n)$ is evaluated according to: $\sum_{i=1}^{n} \pi_i u(x_i)$ where π_i's are nonnegative decision weights which sum to one for all partitions of state space and $u(.)$ is a utility function.

A2: (Monotonicity) First order stochastically dominating acts are preferred to acts which they dominate.

In the original version of Prospect Theory (KT1979), the following assumption was implicitly made:

A3': (Independence of beliefs from tastes) The decision weight π_i depends only on E_i.

To see why we require that decision weights sum to one (see assumption A1) consider the partition E_1, E_2 of the state space and suppose first that $\pi_1 + \pi_2 > 1$. Then it is possible to find an ϵ such that an act $(x_1, E_1; x_1 - \epsilon, E_2)$ is preferred to (x_1, S), given continuity of $U(.)$. This preference however violates monotonicity. Now suppose $\pi_1 + \pi_2 < 1$. Then again it is possible to find an ϵ such that an act (x_1, S) is preferred to $(x_1, E_1; x_1 + \epsilon, E_2)$, although it is stochastically dominated by the latter act. The argument extends straightforwardly to more complex acts. Thus, decision weights should sum to one if we want to sustain monotonicity.

Result 1: Assumptions A1 and A3' imply additivity i.e. for all disjoint events A, B: $\pi_{A \cup B} = \pi_A + \pi_B$.
Proof
By A3' we can define for each event E a decision weight $W(E)$. The decision weights of all events in a given partition of S sum to one by A1. Hence $W(E_1 \cup E_2) = W(S) - W(S \setminus (E_1 \cup E_2)) = W(E_1) + W(E_2)$. Then W is a probability measure and Subjective Expected Utility follows. $\square$

The result above means that we can not implement nonadditive measures, which was a crucial part of Prospect Theory, if we make assumption A3', given the general framework of assumption A1. Therefore we are interested in relaxing assumption A3'. There is some preliminary work before we can do it. First, each act has to be transformed into the rank-ordered act. It suffices to combine equal outcomes together and to reorder them so that an act can be presented as: $(x_1, E_1; ...; x_n, E_n)$; $x_1 < x_2 < ... < x_n$. Define $D_i \equiv E_1 \cup ... \cup E_i$, which describes an event of getting an outcome which is worse or equivalent to E_i. Thus D_i determines the ranking position of an event E_i.

A3: (Rank dependence) The decision weight π_i depends on E_i and D_i.

Given the above assumption, the decision weight of the maximal outcome x_n depends only on E_n, its ranking position being always $D_n = S$. Let's define a function $W(.)$ which will be the decision weight of the highest outcome x_n. It is called a **capacity** and it satisfies the properties of probability measure except for additivity: i.e. $W(\emptyset) = 0$ and $W(S) = 1$. Additionally it satisfies the following requirement: if $A \subset B$ then $W(A) \leq W(B)$.[14]
This last requirement has an intuitive explanation. To see this, consider two acts: $(x, A; y, B \cup C)$ and $(x, A \cup B; y, C)$ where $x > y$. By monotonicity of u: $\pi_{A \cup B} u(x) + \pi_C u(y) \geq \pi_A u(x) + \pi_{B \cup C} u(y)$. Since x is the highest outcome in both acts, we know that: $\pi_{A \cup B} = W(A \cup B)$ and $\pi_A = W(A)$. Because the decision weights on both sides of the above inequality sum to one, we also have: $\pi_{B \cup C} = 1 - W(A)$ and $\pi_C = 1 - W(A \cup B)$. Substituting this into the inequality above and rearranging, we obtain: $(u(x) - u(y))(W(A \cup B) - W(A)) \geq 0$. Since $u(x) > u(y)$, it must be that $W(A \cup B) \geq W(A)$ and because we chose A and B arbitrarily, it follows that: $A \subset (A \cup B) = F$ and for any $A \subset F: \quad W(A) \leq W(F)$.
The above argument shows us the intuitive meaning of the concept of capacity. There is additionally one condition, which we would like to impose on capacity. It is called solvability and it is merely a technical condition, which can be regarded as an equivalent of continuity in case of real valued domains. It states that: $\forall \{A \subset C\} \ \wedge \ \{W(A) \leq p \leq W(C)\} \ \exists B$ s.t. $\{W(B) = p\} \ \wedge \ \{A \subset B \subset C\}$.
Now, that we defined the decision weight for the highest outcome, we want to

[14]If A is a proper subset of B then the inequality is strict.

do it for the other outcomes as well. Consider the following two rank ordered acts:

$$(x_1, E_1; ...; x_i, E_i; x_{i+1}, E_{i+1}; ...; x_n, E_n)$$
$$(x_1, E_1; ...; x_i, E_i; z, (E_{i+1} \cup ... \cup E_n))$$

where $z > x_i$. It is clear from above, that the ranking positions of the first i outcomes are the same for both acts. Also the corresponding outcomes in both acts are contingent on the occurrence of the same events. Hence by A3, those elements have the same decision weights in both acts. Moreover the outcome z in the second act happens to be the highest outcome in this act and hence its decision weight is $W(E_{i+1} \cup ... \cup E_n)$. If we denote the decision weights: π_i^I for the first act and π_i^{II} for the second act, we can write:

$$\begin{aligned} \pi_{i+1}^I + ... + \pi_n^I &= 1 - (\pi_1^I + ... + \pi_i^I) \\ &= 1 - (\pi_1^{II} + ... + \pi_i^{II}) \\ &= W(E_{i+1} \cup ... \cup E_n) \end{aligned}$$

And it follows directly that:

$$\pi_i = \sum_{j=i}^{n} \pi_j - \sum_{j=i+1}^{n} \pi_j = W(E_i \cup ... \cup E_n) - W(E_{i+1} \cup ... \cup E_n) \tag{3}$$

Let's summarize the above argument.

Result 2: Assumptions A1, A2 and A3 imply the so called Choquet expected utility: the rank ordered act $(x_1, E_1; ...; x_n, E_n)$ is evaluated according to: $\sum_{i=1}^{n} \pi_i u(x_i)$, where π_i's are defined by (3) above.

The concept of capacity is quite vague without imposing any further requirements on it. Suppose we want to investigate what restrictions should be imposed on a capacity if a given agent is pessimistic, in the sense that, ceteris paribus, he puts more weight on the events with worse ranking position. Assume that there is an event E yielding outcome x with the ranking position D. Thus, its decision weight is: $W(E \cup D^c) - W(D^c)$. Worsening ranking position means decreasing D. Hence, pessimism implies that if $C \subset D$, then $W(E \cup C^c) - W(C^c) \geq W(E \cup D^c) - W(D^c)$. Define $A = C^c$ and $B = E \cup D^c$. Then notice that:

$$\begin{aligned} A \cup B &= C^c \cup (E \cup D^c) = (C^c \cup E) \cup (C^c \cup D^c) = (C^c \cup E) \cup C^c = E \cup C^c \\ A \cap B &= C^c \cap (E \cup D^c) = (C^c \cap E) \cup (C^c \cap D^c) = \emptyset \cup D^c = D^c \end{aligned}$$

So pessimism implies convex capacity, where **convex capacity** is defined as: $W(A \cup B) + W(A \cap B) \geq W(A) + W(B)$. Similarly optimism implies concave capacity, which occurs when the above inequality is reversed.

The last point in this section concerns a main identifying assumption of rank dependence models, i.e. **comonotonic independence** introduced by Schmeidler (1989). It states that the independence axiom[15] should be obeyed only within comonotonic sets of acts. Comonotonic set of acts consists of acts which have the same ordering of outcomes in terms of events, i.e. there are no states s_i and s_j, such that: $f_i > f_j \wedge g_i < g_j$, for f, g being acts with outcomes f_i, g_i, respectively when state s_i occurs.

Intuitively, since comonotonic acts have rank correlation 1, they cannot be used to hedge away each other's risk through the formation of compound acts. Within comonotonic sets, the decision maker obeys all the Savage axioms locally and hence behaves as Expected Utility maximizer. It suggests that we should use rank dependence models in portfolio management since usually optimal portfolio aims at hedging against risk, which requires operating on different comonotonic sets. In case of the real-valued state space, rank-ordered comonotonic acts correspond to functions which are monotonically nondecreasing in the state space.

Let us stress one more thing. The probability weighting in Prospect Theory[16] implies transforming each probability individually into some associated decision weight. The probability weighting in Rank Dependence Models implies transforming the whole cumulative distribution. Hence, the same value of probability gets different decision weight depending on the ranking position. It is particulary important not to confuse probability distortion function for Prospect Theory with probability distortion function for Rank Dependent models. The difference is especially pronounced for non-simple prospects.[17]

5 Cumulative Prospect Theory

[18]We shall present here the Cumulative Prospect Theory under uncertainty, but we could do similar analysis for the case of risk. As said above, CPT combines the Rank Dependent model with Prospect Theory. In this section we shall use

[15]Recall that the independence axiom means, that preferences between lotteries or acts will be unaffected by substitution of common factors.

[16]We are talking here about choice under risk.

[17]Prospect Theory was originally discussed only with simple prospects with at most two non-zero outcomes.

[18]This section draws on Kahnemann and Tversky (1992).

the term prospect to refer to an act which is defined relative to a reference point. That means, there exists a reference point which is normalized to zero, and all negative outcomes denote losses, and all positive outcomes denote gains. We adopt the same notation as in the previous section. We shall deal with rank-ordered prospects of the following form:

$$f \equiv (x_1, E_1; ...; x_k, E_k; x_{k+1}, E_{k+1}; ...; x_n, E_n) \tag{4}$$

where $x_1 < ... < x_k < 0 < x_{k+1} < ... < x_n$. Let's define a positive and a negative part of f:

$$\begin{aligned} f^+ &\equiv (0, E_1 \cup ... \cup E_k; x_{k+1}, E_{k+1}; ...; x_n, E_n) \\ f^- &\equiv (x_1, E_1; ...; x_k, E_k; 0, E_{k+1} \cup ... \cup E_n) \end{aligned}$$

The property of CPT called **sign dependence**[19] means that we apply different weighting schemes for the negative and for the positive part of a prospect. Negative part is weighted according to $\pi_i^- = W^-(E_1 \cup ... \cup E_i) - W^-(E_1 \cup ... \cup E_{i-1})$, for $i = 1, ..., k$ and positive part is weighted according to $\pi_i^+ = W^+(E_i \cup ... \cup E_n) - W^+(E_{i+1} \cup ... \cup E_n)$ for $i = k+1, ..., n$, where W^- and W^+ are two different nonadditive capacities.[20]

Sign dependence is not just a minor extension implied by reference dependence. To appreciate this fact, notice that in purely positive or purely negative prospects the decision weights necessarily sum to one. We can show this by using equation (3):

$$\begin{aligned} \sum_{i=1}^{n} \pi_i &= \sum_{i=1}^{n} [W(E_i \cup ... \cup E_n) - W(E_{i+1} \cup ... \cup E_n)] \\ &= W(E_1 \cup ... \cup E_n) = W(S) = 1 \end{aligned} \tag{5}$$

However in the case of mixed prospects, Cumulative Prospect Theory does not assume that decision weights should sum to one. With the usual shape of probability weighting function they will rather sum to less than one. This property is called **subcertainty**. To show this we use the above definitions for

[19] In CPT the crucial axiom is sign comonotonic independence, so that independence is satisfied only on the sign comonotonic sets (the same ordering and the same sign).

[20] For the concept of capacity, see the previous section.

decision weights to write:

$$\begin{aligned}\sum_{i=1}^{n}\pi_i &= \sum_{i=1}^{k}\left[W^-(E_1\cup...\cup E_i)-W^-(E_1\cup...\cup E_{i-1})\right]\\ &+ \sum_{i=k+1}^{n}\left[W^+(E_i\cup...\cup E_n)-W^+(E_{i+1}\cup...\cup E_n)\right]\\ &= W^-(E_1\cup...\cup E_k)+W^+(E_{k+1}\cup...\cup E_n)\end{aligned}$$

Recall from the section on intuition of Rank Dependency, that when decision weights do not sum to one, it is possible to construct examples of choice violating monotonicity. However, in case of CPT, even though the decision weights do not necessarily sum to one for mixed prospects, monotonicity is satisfied. The intuitive explanation for this fact is that when constructing examples of non-monotonic behavior we need to compare lotteries with some outcomes changing signs. Where an outcome changes sign, its impact on the CPT representation changes not only via change of weighting but it also has a reversed effect on the CPT representation function.
Having discussed sign dependence, we can now show the CPT representation formula for a given prospect of the form as in (4):

$$V_{CPT}(f)=\sum_{i=1}^{k}\pi_i^-u(x_i)+\sum_{i=k+1}^{n}\pi_i^+u(x_i) \tag{6}$$

Below we present three basic blocks of Cumulative Prospect Theory:

- **prospect theory** - reference dependence, the shape of utility function and weighting function i.e. diminishing sensitivity and loss aversion
- **rank-dependence model** - in case of uncertainty it is called Choquet Expected Utility model (CEU), developed by Schmeidler (1989) and Gilboa (1987), in case of risk it is called rank-dependent model, developed by Quiggin (1982)
- **sign-dependence** - a novel feature allowing for different weighting schemes for gains and losses

There is a vast literature on parametric shape of utility function and weighting function (see for example Prelec (1998), Rieger and Wang (2006) or Kahnemann and Tversky (1992)). We decided not to spend time on this issue, because it is primarily of empirical interest.

6 Risk attitudes in CPT

The big advantage of Expected Utility Theory is its simplicity. The curvature of a von Neumann Morgenstern utility function alone determines risk attitudes of a decision maker. It allows simple characterization and the resulting theory is easily applicable. In case of CPT, risk attitude is characterized by three elements:

- **nonadditive decision weights**
- **loss aversion** - by how much losses loom larger than gains
- **basic utility** - measuring the intrinsic value of particular outcome

In this section we will try to sketch some methods of characterizing risk in CPT setting. We will focus on the first two of the above mentioned elements of risk attitude, because these two elements are novel feature of CPT as compared to EUT.

First, we shall introduce and derive stochastic dominance result for Cumulative Prospect Theory. Before we do it, let us present an extended version of CPT, which allows for continuous outcome space. It is straightforward to show that the CPT utility function in this case is:

$$\begin{aligned} V(x,F) &= \int_{-\infty}^{0} u(x)d[w_-(F(x))] + \int_{\infty}^{0} u(x)d[w_+(1-F(x))] \\ &= \int_{-\infty}^{0} u(x)d[w_-(F(x))] + \int_{0}^{\infty} u(x)d[w_+(F(x))] \end{aligned} \tag{7}$$

where $F(x) = \int_{-\infty}^{x} dp$ is a cumulative distribution function for outcomes. To see how this formulation includes the discrete case, we can set $p(x) = \sum_i \delta_{x_i} p_i$, where δ_x is a Dirac probability mass at x and probabilities satisfy usual requirements.

6.1 Stochastic Dominance in Cumulative Prospect Theory

Stochastic dominance for Cumulative Prospect Theory can be stated as: F is preferred to G iff:

$$\begin{aligned} &\int_{-\infty}^{0} u(x)d[w_-(F(x))] + \int_{0}^{\infty} u(x)d[w_+(F(x))] \\ \geq &\int_{-\infty}^{0} u(x)d[w_-(G(x))] + \int_{0}^{\infty} u(x)d[w_+(G(x))] \end{aligned} \tag{8}$$

Integrating by parts all four integrals in the above inequality results in (a, b are the lower and upper bound for outcome space, they are allowed to be $-\infty$, $+\infty$ respectively):

$$\begin{aligned} & [u(x)w_-(F(x))]_a^0 + [u(x)w_+(F(x))]_0^b \\ & -\int_a^0 u'(x)w_-(F(x))dx - \int_0^b u'(x)w_+(F(x))dx \\ \geq \; & [u(x)w_-(G(x))]_a^0 + [u(x)w_+(G(x))]_0^b \\ & -\int_a^0 u'(x)w_-(G(x))dx - \int_0^b u'(x)w_+(G(x))dx \end{aligned}$$

The first elements both on the LHS and the RHS are zero because $u(0) = 0$, $w_-(F(a)) = 0$ and $w_-(G(a)) = 0$. Both second elements on the RHS and the LHS are equal to $u(b)$ because $u(0) = 0$ and $w_+(F(b)) = 1$ and $w_+(G(b)) = 1$. So they cancel each other. That leaves us with:

$$\int_a^0 u'(x)[w_-(G(x)) - w_-(F(x))]dx + \int_0^b u'(x)[w_+(G(x)) - w_+(F(x))]dx \geq 0$$

Integrating by parts once again gives us the following:

$$\begin{aligned} & \left[u'(x)\int_a^x [w_-(G(u)) - w_-(F(u))]du\right]_a^0 \\ - \; & \int_a^0 u''(x)\int_a^x [w_-(G(u)) - w_-(F(u))]dudx \\ + \; & \left[u'(x)\int_0^x [w_+(G(u)) - w_+(F(u))]du\right]_0^b \\ - \; & \int_0^b u''(x)\int_0^x [w_+(G(u)) - w_+((u))]dudx \geq 0 \end{aligned}$$

And rewriting:

$$\begin{aligned} & u'(0)\int_a^0 [w_-(G(u)) - w_-(F(u))]du \\ - \; & \int_a^0 u''(x)\int_a^x [w_-(G(u)) - w_-(F(u))]dudx \\ + \; & u'(b)\int_0^b [w_+(G(u)) - w_+(F(u))]du \\ - \; & \int_0^b u''(x)\int_0^x [w_+(G(u)) - w_+((u))]dudx \geq 0 \end{aligned} \tag{9}$$

Notice that for $0 < x \leq b$, we have $u''(x) \leq 0$ and hence: if $\int_0^x [w_+(G(u)) - w_+(F(u))]du \geq 0$ holds for all x then the last two terms of the above inequality are nonnegative. On the other hand, for $a \leq x < 0$, we have $u''(x) \geq 0$. Now, let us concentrate on the first two terms of the above inequality, decompose the second term and transform as shown below:

$$\begin{aligned}
& u'(0)\int_a^0 [w_-(G(u)) - w_-(F(u))]du \\
- & \int_a^0 u''(x)\int_a^x [w_-(G(u)) - w_-(F(u))]dudx \\
= & u'(0)\int_a^0 [w_-(G(u)) - w_-(F(u))]du \\
- & \int_a^0 u''(x)\int_a^0 [w_-(G(u)) - w_-(F(u))]dudx \\
+ & \int_a^0 u''(x)\int_x^0 [w_-(G(u)) - w_-(F(u))]dudx \qquad (10)
\end{aligned}$$

Notice that in the second term above we can now separate the two integrals:

$$\begin{aligned}
& -\int_a^0 u''(x)dx\int_a^0 [w_-(G(u)) - w_-(F(u))]du \\
= & -u'(0)\int_a^0 [w_-(G(u)) - w_-(F(u))]du + u'(a)\int_a^0 [w_-(G(u)) - w_-(F(u))]du
\end{aligned}$$

We can observe that the first element on the RHS of the above equation cancels with the first element on the RHS of equation (10). Going back to the whole inequality (9), we can write:

$$\begin{aligned}
& u'(a)\int_a^0 [w_-(G(u)) - w_-(F(u))]du \\
+ & \int_a^0 u''(x)\int_x^0 [w_-(G(u)) - w_-(F(u))]dudx \\
+ & u'(b)\int_0^b [w_+(G(u)) - w_+(F(u))]du \\
- & \int_0^b u''(x)\int_0^x [w_+(G(u)) - w_+((u))]dudx \geq 0
\end{aligned}$$

The above shows now that: if $x \geq 0$ then $u''(x) \leq 0$ and we require $\int_x^0 [w_-(G(u)) - w_-(F(u))]du = \int_0^x [-(w_-(G(u)) - w_-(F(u)))]du \leq 0$ and if $x \leq 0$ then $u''(x) \geq 0$ and we require $\int_x^0 [w_-(G(u)) - w_-(F(u))]du \geq 0$.

The above derivation may seem a bit messy, so let me summarize. We have just demonstrated that:

$$\begin{aligned}
\forall x \qquad & \int_0^x [w_+(G(u)) - w_+(F(u))]du \geq 0 \;\text{ and} \\
& \int_x^0 [w_-(G(u)) - w_-(F(u))]du \geq 0 \qquad (11)
\end{aligned}$$

if and only if F dominates G by stochastic dominance in CPT as defined in the beginning of this section.[21] It is possible to design examples in which neither F nor G dominates the other by First Degree Stochastic Dominance (FSD), Second Degree Stochastic Dominance (SSD) and still there is a stochastic dominance in the above sense, even in the case in which $w_-(.)$ and $w_+(.)$ are identity functions.[22]

6.2 Probability weighting

[23]In this subsection we focus attention on the probability weighting for gains, because the analysis for losses is identical. So we suppress the superscript "+". We want to formalize the fact that the probability distortion function has an inverse-S shape. We will however concentrate on the case of uncertainty where there is actually no probability distortion function, because there is no given probability. But we can always regard a capacity as nonlinear distortion of a subjective probability. Needless to say, modeling risk attitudes in uncertainty case is very similar conceptually to modeling risk attitudes under risk, except for the fact that uncertainty case is more general since it does not assume the knowledge of objective probabilities.

A capacity W satisfies **subadditivity** (SA), if there are events E, E' such that:

$$W(B) \geq W(A \cup B) - W(A) \quad \text{whenever} \quad W(A \cup B) \leq W(S - E) \tag{12}$$

$$1 - W(S - B) \geq W(A \cup B) - W(A) \quad \text{whenever} \quad W(A) \geq W(E') \tag{13}$$

The condition (12) is called **lower SA** and the condition (13) is called **upper SA**. The events E, E' are called lower and upper boundary events. These are "small" events, independent of A and B. For future purposes, let's define $A \succcurlyeq B$ if there exists a gain y such that $(y, A) \succcurlyeq (y, B)$. Obviously: $A \succcurlyeq B$ iff $W(A) \geq W(B)$. We want to show below that these conditions imply the observed preference conditions. Recall that in Prospect Theory discussed in one of the previous sections, there were two effects which suggested nonlinear distortions of probability: certainty effect and turning highly unlikely possibility

[21]Well, we showed only the implication in one direction, however the implication in the other direction follows a similar procedure and hence, is omitted here.

[22]No probability distortion case.

[23]This subsection partly follows Tversky and Wakker (1995).

into impossibility. Formally, we observe **certainty effect** if: [24]

$$(x, S - B) \sim (y, A) \Rightarrow (x) \succcurlyeq (y, A; x, B), \quad \text{where } 0 < x < y, A \succcurlyeq E' \qquad (14)$$

The above statement can easily be derived from the observed choice characteristics in Allais paradox and from continuity and monotonicity. Observe that we can obtain the acts on the RHS of the above implication by changing: B causing 0 to B causing x in the acts on the LHS. Note that outcome x in the left lottery on the LHS was uncertain and in the left lottery on the RHS it became certain. Rewrite (14) in terms of CPT:

$$\begin{aligned} u(x)W(S-B) &= u(y)W(A) \Longrightarrow \\ \Longrightarrow u(x) &\geq u(y)W(A) + u(x)(W(A \cup B) - W(A)) \end{aligned} \qquad (15)$$

Now assume, that upper SA holds and multiply both sides of (13) by $u(x)$ and add and subtract $u(y)W(A)$ from the RHS. We obtain:

$$u(x)(1 - W(S-B)) \geq u(x)(W(A \cup B) - W(A)) + u(y)W(A) - u(y)W(A)$$

Now we substitute $u(x)W(S-B) = u(y)W(A)$ from (15) into above inequality and rearrange:

$$\begin{aligned} u(x) - u(x)W(S-B)) &\geq u(x)(W(A \cup B) - W(A)) + u(y)W(A) - u(x)W(S-B) \\ u(x) &\geq u(y)W(A) + u(x)(W(A \cup B) - W(A)) \end{aligned}$$

Hence, we showed that upper SA implies certainty effect in CPT. Tversky and Wakker (1995) show the implication in the other direction as well.
The other effect leading to nonadditive distortions of subjective probability is **turning impossibility into possibility**. This leads to overweighting of small probabilities of extreme events. Formally:

$$\begin{aligned} (x) \sim (y, A; x, B) &\Rightarrow (y, B; x, S - B) \succcurlyeq (y, A \cup B) \\ &\text{where } \ 0 < x < y, \ A \cup B \preccurlyeq S - E \end{aligned} \qquad (16)$$

The above statement can again be easily derived from the observed choice characteristics and from continuity and monotonicity. Note that we can obtain the acts on the RHS of the above implication by changing: B causing x to B causing y in the acts on the LHS. This means that outcome y was impossible

[24] We denote (x) as an act yielding x with certainty (S) and (x, A), where A is a proper subset of S, as an act yielding x contingent on the occurrence of A and zero otherwise (event $S - A$). More complex acts we denote in a similar way.

in the left lottery on the LHS of the above implication and became possible in the left lottery on the RHS of the above implication. Rewrite (16) in terms of CPT:

$$\begin{aligned} u(x) &= u(y)W(A) + u(x)(W(A \cup B) - W(A)) \Longrightarrow \\ \Longrightarrow u(y)W(A \cup B) &\leq u(y)W(B) + u(x)(1 - W(B)) \end{aligned} \tag{17}$$

Now assume lower SA holds and multiply both sides of (12) by $u(y) - u(x)$ and substitute $u(x) = u(y)W(A) + u(x)(W(A \cup B) - W(A))$ from (17), or $-u(x)(W(A \cup B) - W(A)) = u(y)W(A) - u(x)$ into the resulting inequality. We obtain then:

$$\begin{aligned} (u(y) - u(x))W(B) &\geq u(y)(W(A \cup B) - W(A)) + u(y)W(A) - u(x) \\ u(y)W(B) + u(x)(1 - W(B)) &\geq u(y)W(A \cup B) \end{aligned}$$

And hence we showed that lower SA implies the effect of overweighting small probabilities. Again, Tversky and Wakker (1995) proved also the implication in the other direction. To summarize:

Result on subadditivity: Under the usual requirements, the weighting function W satisfies SA iff (14) and (16) are satisfied.

It should be emphasized here, that lower and upper subadditivity should be interpreted with caution. The motivation for introducing these conditions was the observed pattern of choices - paying to much attention to extreme events and too little attention to intermediate events. Suppose we switch for the moment to the risk situation and imagine we have a probability distortion function $w(.)$ which transforms cumulative probabilities. Lower subadditivity for risk can be written as: $w(p) \geq w(p+q) - w(q), \;\; \text{for} \;\; w(p+q) \leq w(1-\epsilon)$, where ϵ is a boundary probability. We can transform this condition into: $\frac{w(p)-w(0)}{p} \geq \frac{w(p+q)-w(q)}{p}$ and letting p approach zero we obtain: $w'(0) \geq w'(q), \;\; \text{for} \;\; w(q) \leq w(1-\epsilon)$. So the function $w(.)$ is concave for probabilities in the interval $[0, 1-\epsilon]$. Similarly, upper SA for risk can be written as: $1 - w(1-p) \geq w(p+q) - w(q), \;\; \text{for} \;\; w(q) \geq w(\epsilon')$, where ϵ' is again a boundary probability. Transforming this condition results in: $\frac{w(1)-w(1-p)}{p} \geq \frac{w(p+q)-w(q)}{p}$ and letting p approach zero we obtain: $w'(1) \geq w'(q), \;\; \text{for} \;\; w(q) \geq w(\epsilon')$. Thus the function $w(.)$ is convex for probabilities in the interval $[\epsilon', 1]$. It is harmless to assume that $\epsilon' \leq 1 - \epsilon$. In this case, we have a concave region for probabilities in $[0, \epsilon']$, possibly linear region for probabilities $[\epsilon', 1-\epsilon]$,[25] and a convex region for probabilities $[1-\epsilon, 1]$. To

[25]Only if $\epsilon' < 1 - \epsilon$.

sum up:

$$w''(p) \begin{cases} \leq 0 & \text{for} \quad p \in [0, \epsilon'] \\ = 0 & \text{for} \quad p \in [\epsilon', 1-\epsilon] \\ \geq 0 & \text{for} \quad p \in [1-\epsilon, 1] \end{cases} \tag{18}$$

For this kind of function it can happen, that it doesn't have a fixed point in the interior of $[0, 1]$ interval. It is possible in two cases: either $\lim_{p \to 0} w'(p) < 1$ or $\lim_{p \to 1} w'(p) < 1$. In the first case, we observe extreme overweighting of small probabilities of high ranked events and no overweighting of small probabilities of low ranked events ($w(p)$ lies entirely below 45-degree line for the interior of $[0, 1]$). In the second case, we observe extreme overweighting of small probabilities of low ranked events and no overweighting of small probabilities of high ranked events ($w(p)$ lies entirely above 45-degree line for the interior of $[0, 1]$). There are at least two important implications of the above demonstrations:

First, if we want the probability weighting function to exhibit overweighting of both small high ranked and and small low ranked events, we have to impose additional condition on a weighting function, which will ensure the existence of a fixed point in the interior of $[0, 1]$ interval. We can simply do it by requiring $\lim_{p \to 0} w'(p) > 1$ and $\lim_{p \to 1} w'(p) > 1$. Together with lower and upper SA, it guarantees the existence of a fixed point in the interior of $[0, 1]$. This will fix a problem in situations under risk. Additionally we can define an index of lower (upper) SA as: $\upsilon_{LSA} \equiv \left[\int_0^{p*}(w(p) - p)dp\right]^+$ ($\upsilon_{USA} \equiv \left[\int_{p*}^1 (p - w(p))dp\right]^+$), where $p*$ is a fixed point of $w(p)$.

Second, the possibility of cases such as described above[26] suggests a new way of looking at pessimism and optimism. In some of the earlier sections, we introduced a concept of pessimism (optimism), which was shown to imply convex (concave) capacity. In case of risk it is easy to show that pessimism (optimism) implies concave[27] (convex) weighting function. The above anomalies of weighting function suggest however that concave (convex) weighting function does not imply pessimism (optimism), and hence the implication can be shown only in one direction. In particular, it is easy to design an example in which weighting function is concave almost on the whole domain, but there is no overweighting of lower ranked events and extreme overweighting of high ranked events.[28] This suggests that we should distinguish local pessimism/optimism

[26] I.e. no overweighting of small high ranked (low ranked) events and extreme overweighting of small low ranked (high ranked) events.

[27] Attention: the equivalent of **convex** capacity in case of risk is **concave** probability weighting function.

[28] The same for convex weighting function

from global pessimism/optimism. Local pessimism (optimism) implies concavity (convexity) of a weighting function. Global pessimism (optimism) would imply that $\forall p: \ w(p) \geq p$ ($\forall p: \ w(p) \leq p$). In situations, in which we have a mix of pessimism (for low values of p) and optimism (for high values of p), we are assured that a fixed point of $w(p)$ exists in the interior of $[0,1]$. Then we can use the above introduced indexes υ_{LSA} and υ_{USA} to measure the degree of optimism within the optimistic part and the degree of pessimism within the pessimistic part. To understand better the probability weighting in CPT, especially the difference between probability weighting with small number of outcomes and probability weighting with continuous outcomes, it might be useful to consider function $w'(p)$ instead of $w(p)$. The weight of a particular event would then be determined according to: $\pi_i = \int_{1-F(x_i)}^{1-F(x_{i-1})} w'(p)dp$ (for losses $\pi_i = \int_{F(x_{i-1})}^{F(x_i)} w'(p)dp$ accordingly). In case of continuous outcomes, $F(x_{i-1})$ would be arbitrarily close to $F(x_i)$ and so the decision weight in this case would be just $\pi(x) = w'(F(x))$. Moreover in the continuous time, it is easier to characterize which probabilities are overweighted and which are underweighted. All p for which $w'(p) > 1$, are overweighted and all p for which $w'(p) < 1$ are underweighted. This implies low ranked and high ranked events are overweighted and moderately ranked events are underweighted.

After this extensive discussion on the issue of subadditivity, we want to present results concerning comparative subadditivity and as before, we do it for the case of uncertainty. A transformation $\tau : [0,1] \rightarrow [0,1]$ is called SA if it satisfies the same requirements as $W(.)$ does in (12) and (13). A weighting function W_2 is more SA than W_1, if it is obtained from W_1 by SA transformation. We will prove below the necessity part of the following equivalence:

Result on comparative SA: W_2 is more SA than W_1 iff W_2 is a strictly increasing transform of W_1 and:

$$W_1(C) = W_1(A \cup B) - W_1(A) \Rightarrow W_2(C) \geq W_2(A \cup B) - W_2(A) \qquad (19)$$

$$\wedge$$

$$1 - W_1(S-C) = W_1(A \cup B) - W_1(A) \Longrightarrow$$
$$\Longrightarrow 1 - W_2(S-C) \geq W_2(A \cup B) - W_2(A) \qquad (20)$$

with boundary condition for lower comparative SA (19): $W_1(A \cup B) \leq W_1(S - E)$ for some E and for upper comparative SA (20): $W_1(A) \geq W_1(E')$ for some E'.

Proof ($\Rightarrow$): To prove the necessity of this result we just assume $W_1(C) = W_1(A \cup B) - W_1(A)$ and write $W_2(C) = \tau(W_1(C)) \geq \tau(W_1(C) + W_1(A)) - \tau(W_1(A))$ and $W_2(C) = \tau(W_1(C)) \geq \tau(W_1(A \cup B)) - \tau(W_1(A)) = W_2(A \cup B) - W_2(A)$, by using our assumption and the lower SA property of τ. The same for comparative upper SA: assume $1 - W_1(S - C) = W_1(A \cup B) - W_1(A)$ and write $1 - W_2(S - C) = 1 - \tau(W_1(S - C)) \geq \tau(W_1(A \cup B)) - \tau(W_1(A))$ and $1 - W_2(S - C) \geq W_2(A \cup B) - W_2(A)$, by using the upper SA property of τ and our assumption. This proves the necessity part of the above result. $\square$

We could also state the corresponding preference conditions for the proposition that $\succcurlyeq_2$ is more SA than $\succcurlyeq_1$. However the statement follows very similar lines as the statements concerning the analysis of subadditivity (the relation between SA on the one side, certainty effect and overweighting of extreme events probabilities on the other side), and hence we omit it here. Another important thing which we omit here is a so called source sensitivity which measures the preference over sources of uncertainty. This and other related results can be found in Tversky and Wakker (1995). The important empirical finding which we should underscore here is that uncertainty enhances the departures from expected utility as compared to risk. The nonadditivity in weighting schemes under uncertainty is more pronounced than nonlinearity in weighting schemes under risk. Moreover it is found that people prefer risk to uncertainty when they feel incompetent. In other situations, when they don't feel incompetent or ignorant about the subject, people often prefer to bet on an uncertain source.

6.3 Loss aversion

[29]We assume here that there exists a basic utility function U that reflects the intrinsic value of outcomes for the individual. Because of the psychological perception of a reference point, however people evaluate losses differently than gains. The overall utility u is a composition of a loss aversion index $\lambda > 0$, and the basic utility U. That means we can write:

$$u(x) = \begin{cases} U(x) & \text{if } x \geq 0 \\ \lambda U(x) & \text{if } x < 0 \end{cases} \tag{21}$$

Now, we can assume that the basic utility U is smooth everywhere but particularly at the reference point, and the only reason, the kink appears in u at the reference point is because of $\lambda > 1$ (losses loom larger than gains). This is

[29]This subsection partly follows Koebberling and Wakker (2005).

obviously quite strict assumption but it reflects the psychological importance of a reference point and enables disentangling basic utility and loss aversion parts of risk aversion. Furthermore, the empirical findings suggest that basic utility embodies the intrinsic value of outcomes, whereas loss aversion and probability weighting is psychological in nature. Koebberling and Wakker (2005) propose the following loss aversion index: $\lambda = \frac{\lim_{x\to 0} u(-|x|)}{\lim_{x\to 0} u(|x|)}$, if the limits $\lim_{x\to 0} u(-|x|)$ and $\lim_{x\to 0} u(|x|)$ exist. We immediately see that using index λ, requires assuming that loss aversion is a constant fraction of basic utility. The clear advantage of this approach is that it is independent of the unit of payment. Changes in scale do not affect the value of this index. The implicit scaling convention in this definition is that the function U is smooth at zero, so that the left and right derivative of this function agree at the reference point. We can adopt different scaling convention, i.e. we can choose $y > 0$ and set $-U(-y) = U(y)$ which then results in an index $\lambda_1 = \frac{-u(-y)}{u(y)}$. However such scaling conventions are not independent of the unit of payment anymore. They thus change under different scaling of outcomes.

The above loss aversion index assumes that it is possible to separate basic utility from loss aversion. Kahnemann and Tversky (1979) define loss aversion in the following way:

$$\text{for } y > x \geq 0: \quad u(x) + u(-x) > u(y) + u(-y) \tag{22}$$

From this definition we can derive two other conditions. First, if we set $x = 0$ then $u(y) < -u(-y)$. This suggests the above y-scaling condition with the loss aversion index $\lambda_1 = \frac{-u(-y)}{u(y)}$. Second, we can let y approach x. Defining $y = x+\Delta$, we have from (22): for $x \geq 0, \Delta > 0$: $\frac{u(-x)-u(-x-\Delta)}{\Delta} > \frac{u(x+\Delta)-u(x)}{\Delta}$, and letting Δ approach zero, we obtain: $u'(-x) > u'(x)$, so that the utility function is steeper for losses than for gains. The condition: $\forall x > 0: \ u'(-x) > u'(x)$ is obviously stronger than: $\forall x > 0: -u(-x) > u(x)$. It suggests that maybe we should define another index - the index of local loss aversion: $\lambda_2 = \frac{-u'(-x)}{u'(x)}$. This index will inform us, how much more an individual dislikes an additional marginal loss, given loss of x than he likes an additional marginal gain, given gain of x. Local loss aversion index, however, would not be separable from basic utility, which is the main advantage of the Koebberling and Wakker (2005) formulation (index λ). To sum up, global index of loss aversion λ is useful because it separates basic utility from loss aversion, but the implied concept of loss aversion in this setting is weaker than assumed by Kahnemann and Tversky (1979, p.279).

7 Cumulative Prospect Theory - Applications and Concluding Thoughts

[30]There is a huge literature demonstrating that the Expected Utility paradigm works pretty well in a wide variety of situations. Why should we then bother to search for some other theories, such as CPT, which are certainly more complex and more difficult to apply? Let me give you just a few prominent examples which answer this question directly:

- **Equity premium puzzle** - The average observed return to stocks is higher approximately by 8 per cent than bond returns. Mehra and Prescott (1985) showed that under the standard assumptions of EUT, investors must be extremely risk averse to demand such high a premium, which created a puzzle. Benartzi and Thaler (1995) suggested an answer based on reference dependence, crucial aspect of CPT. They argued, that in the short run, for example annually, stock returns are negative much more frequently than bond returns. Loss averse investors will then naturally demand large equity premium to compensate for the much higher chance of losing money.
- **Disposition effect** - Investors are observed to hold on to stocks that have lost their value, compared to their purchase price, too long and are eager to sell stocks that have risen in value too soon. It suggests that investors are willing to gamble in the domain of losses and are risk averse in the domain of gains, exactly as predicted by reference dependence. Expected Utility rules on the other hand, would advise you to keep the stocks as long as you expect them to grow, and sell them, as long as you expect them to fall, irrespective of the purchase price.
- **Permanent income hypothesis** - According to this classic hypothesis, people should anticipate their lifetime income and spend the constant fraction of it every period. However, the observed behavior is different. In particular, it is commonly observed that people spend more, when their future wages are expected to increase, but they do not cut back when their future wages are cut. A perfectly suitable explanation would be, that: first, loss aversion makes people feel awful, when they cut consumption; second, due to reflection effect, people are willing to gamble, that next year's wages may turn out to be better after all.

[30]This section partly follows Camerer (1998).

- **Racetrack betting, state lotteries, insurance** - The nonlinear weighting of probabilities is capable of explaining a lot of observed behavior coming from different situations. In racetrack betting, people tend to commonly overbet longshots - horses with relatively small chance of winning. In case of state lotteries it was observed that large cumulated jackpots attract huge number of people. In terms of Expected Utility, it can only be explained by a utility function, which is convex in money. In case of insurance, people often buy insurance against very small risks. In standard Expected Utility, a person who is averse to a tiny risk should be more averse to big risks. Rabin (2000) was the first, who demonstrated how dramatic the implications of local risk aversion are for global risk aversion. Hence the aversion for tiny risks would result in enormous aversion for bigger risks, if we were to stick to EUT. All these phenomena, and these are just few examples, can be explained by nonlinear weighting of probabilities, in particular by overweighting of small probabilities of extreme outcomes.

The above examples merely give a touch of flavor of how powerful in explaining real world phenomena CPT can be. Loss aversion, reflection effects[31] and non-additive weighting are key features of CPT and to appreciate them we need to enter the world in which EUT sees only paradoxes or puzzles. To summarize, I will try to sketch the most important situations, in which we can expect to be better off by applying CPT instead of EUT.

First, we need to have an environment, in which it is reasonable to assume that people are isolating or bracketing the relevant decisions. Otherwise, the reference point is difficult to define.

Second, the departure from expected utility due to nonlinear weighting shall be particularly strong in the presence of some extreme events happening with non-negligible probability. Non-negligible, because people overweight small probabilities of extreme events, provided that they notice them. If probabilities are too small people are likely to neglect them. The default probability of one firm is likely to be non-negligible, but the probability of a major market crash is likely to be negligible in most situations. So distributions with heavy tails, skewed distributions are likely to produce larger departures from Expected Utility Theory. Distributions can be skewed in a usual sense and also skewed relative to the reference point - more probability mass put on losses than on gains or the opposite. Situations like modeling default, insurance or even usual portfolio

[31] In particular gambling in the domain of losses.

management commonly involve these kinds of distributions.
Third, departures from classical theory can be expected for situations in which people perceive some outcomes as losses. Recall, that the utility function for losses is convex and hence people are likely to gamble in the domain of losses, contrary to EUT. Also, situations which involve constant shifts of reference are likely to generate differences between CPT and EUT predictions, because these shifts change the gain/loss status of outcomes.
Fourth, we should expect larger departures from EUT for situations involving uncertainty rather than risk.[32] The additional issue is also the degree to which decision makers feel comfortable or familiar with a given choice situation. If they feel ignorant, they are likely to produce bigger deviation from EUT. The same argument implies that people like professional market traders should violate EUT less often.
The above listing consists of some loose thoughts about the range of applications for CPT. I believe that future research will provide the constantly improving answer to this question. Many topics in finance, insurance and also in economics await being modeled via CPT. There is certainly a lot to be learned from this modeling. Even proving that some classic results are robust to a change from EUT to CPT provides deeper understanding on the importance of different assumptions underlying the theory. It is however certain that many classic results are not robust to a change from EUT to CPT, and hence they need reevaluation. It is hoped that this article demonstrated how an intuitive idea of Cumulative Prospect Theory evolved from experimental and theoretical literature and more importantly how it can be applied in modeling situations under risk and uncertainty.

References

Benartzi, S. and R. Thaler (1995). Myopic loss aversion and the equity premium puzzle. *Quarterly Journal of Economics 110*, 73–92.

Camerer, C. F. (1998, May). Prospect theory in the wild: Evidence from the field. Social Science Working Paper 1037 California Institute of Technology.

Diecidue, E. and P. Wakker (2001). On the intuition of rank-dependent utility. *Journal of Risk and Uncertainty 23*(3), 281–298.

[32] In real life, what we actually always face is uncertainty, not risk. On the other hand, there are uncertain situations which are closer to or further from risk. In some cases, for example, we can estimate probabilities of events and use them as though we were facing risk.

Ellsberg, D. (1961). Risk ambiguity, and the savage axioms. *Quarterly Journal of Economics 75*, 643–669.

Gilboa, I. (1987). Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics 16*, 65–88.

Kahnemann, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–292.

Kahnemann, D. and A. Tversky (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 297–323.

Koebberling, V. and P. Wakker (2005). An index of loss aversion. *Journal of Economic Theory 122*, 119–131.

Machina, M. J. (1983). The economic theory of individual behavior toward risk: Theory, evidence, and new directions. *Technical Report 433, Department of Economics, Stanford University.*

Machina, M. J. (2005). *The New Palgrave Dictionary of Economics* (2nd ed.)., Chapter Non-Expected Utility Theory. Palgrave Macmillan (Basingstoke and New York).

Mehra, R. and E. Prescott (1985). The equity premium puzzle. *Journal of Monetary Economics.*

Nau, R. (2004). Ph.d. seminar on choice theory. notes and readings for class no. 6.

Neumann, v. J. and O. Morgenstern (1947). *The Theory of Games and Economic Behaviour* (2 ed.). Princeton University Press.

Prelec, D. (1998). The probability weighting function. *Econometrica 66*(3), 497–527.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization 3*, 323–343.

Rabin, M. (2000). Risk aversion and expected-utility theory: A calibration theorem. *Econometrica 68*(5), 1281–1292.

Rieger, M. O. and M. Wang (2006). Cumulative prospect theory and the st. petersburg paradox. *Economic Theory 28*, 665–679.

Savage, L. J. (1954). *The Foundations of Statistics*. NY: Wiley.

Schmeidler, D. (1989). Subjective probability and expected utlity without additivity. *Econometrica 57*, 571–597.

Starmer, C. (2000). Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature 38*(2), 332–382.

Tversky, A. and P. Wakker (1995). Risk attitudes and decision weights. *Econometrica 63*(6), 1255–1280.