Prospect Theory Versus Expected Utility Theory: Assumptions, Predictions, Intuition and Modelling of Risk Attitudes

Michał Lewandowski*

Abstract

The main focus of this tutorial/review is on presenting Prospect Theory in the context of the still ongoing debate between the behavioral (mainly descriptive) and the classical (mainly normative) approach in decision theory under risk and uncertainty. The goal is to discuss Prospect Theory vs. Expected Utility in a comparative way. We discuss: a) which assumptions (implicit and explicit) of the classical theory are being questioned in Prospect Theory; b) how does the theory incorporate robust experimental evidence, striving, at the same time, to find the right balance between the basic rationality postulates of Expected Utility (e.g. monotonicity wrt. First-Order Stochastic Dominance), psychological plausibility and mathematical elegance; c) how are risk attitudes modeled in the theory. In particular we discuss prospect stochastic dominance and the three-pillar structure of modeling risk attitudes in Prospect Theory involving: the non-additive decision weights with lower and upper subadditivity and their relationship to the notions of pessimism and optimism, as well as preferences towards consequences separated into preferences within and across the domains of gains and losses (corresponding to basic utility and loss aversion), d) example applications of Prospect Theory.

Keywords: rank-dependence, independence, loss aversion, prospect stochastic dominance, pessimism and optimism

JEL Classification: D81, D03

*Warsaw School of Economics; e-mail: michal.lewandowski@sgh.waw.pl

Michał Lewandowski

1   Introduction

Experimental evidence against the standard Expected Utility EU, in short) model has been accumulated and discussed for at least half-a-century (e.g. Allais, 1953, Rabin, 2000). Out of many behavioral theories proposed in order to accommodate this evidence, the prevailing one is Prospect Theory (PT in short, Kahneman, Tversky, 1979, Tversky, Kahneman, 1992) that combines important ideas from psychology, economics and mathematics into an eclectic model of decision-making in the context of both risk and uncertainty. There are many papers that analyze, describe, apply or review Prospect Theory. This paper belongs to the last category. There are at least several good recent reviews of Prospect Theory, including the book by Wakker (2010) and the survey by Barberis (2013). Why do we need another review?

   The focus of this review is to look at PT from the comparative perspective, by relating important elements of this theory to its main competitor - EU Theory. We concentrate on the principles and assumptions underlying the two theories and discuss the way risk attitudes are modeled in the two theories. By setting PT against EU Theory, we necessarily engage in the ongoing and yet non-settled debate between the behavioral economics supporters on one hand (with the prevailing descriptive focus), and the representatives of the classical/rational economics (with the prevailing normative focus). Our view in this debate is the following: we are neither supporters nor deniers of PT. We do not share the view of Rabin, Thaler (2001) that EU is an Ex-hypothesis and a dead parrot (alluding to the famous sketch of Monty Python) and that it should be replaced by PT. EU sets the standard of rationality in the context of decision-making under risk. Therefore, it has proved very strong in giving correct predictions in many areas of its applications. It is also true that there are contexts in which people consistently violate the axioms of EU and that PT can accommodate most of these violations. However, it does not imply that it is superior to EU. In fact, Barberis (2013) wonders why after 30 years of existence, PT does not have nearly as many as a tiny fraction of applications as compared to the EU models. The reason might be the following: PT has more degrees of freedom, so while it may accommodate more evidence, it gives much weaker testable predictions: instead of a single utility function as in EU, it has a number of free elements, namely a reference dependent value function, in which the reference point is also a free parameter and the two sign-dependent probability weighting functions. It has also been shown that there are alternative approaches that accommodate the same evidence as PT and yet differ significantly on the explanation (See for example Kontek, Lewandowski, 2017). It seems that a good strategy is to seek for a model that achieves proper balance between testability, consistency and normative appeal on one hand, and descriptive accuracy and psychological plausibility on the other. In other words, a good model should explain robust experimental evidence and depart as little from rationality as possible at the same time. The author's view is that there are some elements of PT that fit into this scheme and others that do not. For example, there are models based on reference dependence that are shown to accommodate many EU paradoxes and

yet do not require probability weighting and hence adhere to the EU hypothesis (See Lewandowski, 2017 or Schneider, Day, 2016). To sum up, our strategy here is to present Prospect Theory without avoiding the critical issues.

Other existing reviews are different both in scope and focus and are perhaps less critical than the current one. Apart from the size difference, the book by Wakker (2010) is a textbook treatment focused on demonstrating the link between formal Prospect Theory and its assumptions on one hand and its empirical content and meaning on the other hand. It is also focused on understanding but on a different, more operational, level. It literally digs in the axioms and shows how they are related to the design and conduct of experiments used to elicit preferences. The review of Barberis (2013), on the other hand, focuses on examining difficulties in applying Prospect Theory and tries to encourage more research in this area. It discusses some recent applied work and points to the directions that are potentially worth studying. It does not directly speak to the problem of the descriptive nature of Prospect Theory. By having many more degrees of freedom than the standard Expected Utility model, Prospect Theory cannot have comparably strong predictions unless some additional strong assumptions are imposed that replace the relaxed assumptions of the standard Expected Utility model. The current review addresses this issue. Apart from a pure review of existing results, current paper contains original discussion on prospect stochastic dominance and the way pessimism and optimism can be modeled in CPT.

This paper is organized as follows. Section 2 describes Expected Utility Theory (EUT for short) and its main (implicit and explicit) assumptions. Section 3 reviews the evidence against the standard EU model. Section 4 discusses the original version of Prospect Theory (PT for short) for decision-making under risk and the main problem with it that led to the development of a new, improved version, of the theory. Section 5 presents the intuition behind rank-dependence in the general context of decision-making under uncertainty. Section 6 presents Cumulative Prospect Theory (CPT for short) for decision-making under risk and uncertainty. Section 7 discusses the way attitudes towards risk and uncertainty are incorporated in CPT. This section contains the author's original contribution on prospect stochastic dominance and the pessimism and optimism attitudes. Section 8 presents examples of CPT applications and concludes.

2 Expected Utility Theory

We shall now discuss Expected Utility Theory and its main postulates. We start with properties that are hidden behind the modeling structure.

2.1 Implicit assumptions behind the modeling structure

The first fundamental postulate of decision theory in general is *consequentialism*, stating that choices are judged solely by their consequences. Uncertainty is a situation

Michał Lewandowski

in which the consequence depends on a state of nature that occurs. Let S be an exhaustive mutually exclusive set of states of nature. Acts are finite-valued functions $f : S \to X$ where $X \subset \mathbf{R}$ is an outcome space. Decisions under uncertainty are modeled by assuming preferences over acts. The set of all acts is denoted by A.

This modeling choice hides some assumptions: first, the state space and consequences are known; second, we can assign consequences to each of the states; third, since the decision maker chooses among acts, all that matters in decision depends on S, X and the assignment of one to another. However, the real limitation imposed by these assumptions depends on how S and X are defined. For example, instead of assuming a given state space we may define them as functions from acts to outcomes in order to reflect possible causal relationships between the decision maker's actions and consequences. Alternatively, states could reflect the protocol by which the decision maker obtained information about possible consequences and actions (See Gilboa, 2009, chapter 12).

Risk is a special case of uncertainty in which probability distribution over states of nature is objectively given and known to the decision maker. This distinction between uncertainty and risk is due to Knight (1921). Machina, Schmeidler (1992) propose that the distinction should be between objective (risk) and subjective uncertainty. What is meant by objectively given probability should be understood the following way (this distinction is due to Anscombe, Aumann, 1963): we call a lottery a roulette wheel lottery if it is equivalent to an experiment that may be replicated independently many times with frequency of different possible events converging to some common number (objective probability) as the number of replications increases – in accordance with the law of large numbers. We call a lottery a horse-race lottery if an experiment such as above cannot be defined. In such case we define subjective probability (de Finetti, 1937) of a subset A of S as the price set by the decision maker for the bet paying 1 if A occurs and 0 otherwise, such that the decision maker is indifferent between buying and (short) selling the bet at this price.

Let π be a (objective) probability measure that is defined on all subsets of S. A probability distribution $P : X \to [0,1]$ induced by an act f is defined as (the state space S might be either finite or infinite):

$$P(x) = \int_{\{s \in S : f(s) = x\}} f(s) \tag{1}$$

for $x \in \text{Img}(f)$ and $P(x) = 0$, otherwise. Let $L(X)$ be the set of all such finite-support probability distributions. Decisions under risk are modeled by assuming preference over $L(X)$. It is assumed that this set is endowed with a mixing operation, i.e. if P, Q are two elements of $L(X)$ and α is some number in the interval $[0,1]$, then $\alpha P + (1 - \alpha) Q$ also belongs to $L(X)$ and the following holds:

$$(\alpha P + (1 - \alpha) Q)(x) = \alpha P(x) + (1 - \alpha) Q(x), \text{ for } x \in X \tag{2}$$

This modeling choice hides at least two additional assumptions besides the assumptions implicit in the context of uncertainty and the modeling choice of acts:

First, compound lotteries such as $\alpha P + (1 - \alpha)Q$ are equivalent to appropriately reduced simple lotteries via the above mixing definition (see 2). This is sometimes explicitly stated as the axiom of reduction of compound lotteries. It abstracts from all the joy in the gambling process: it does not matter whether the uncertainty is resolved in one or many stages. Second, this formulation precludes taking into account dependence between different choice alternatives. As shown above (see equation 1) the formulation under risk disposes of the state space, in which dependence is naturally present.

2.2 Expected Utility axioms

Having stated the implicit assumptions present both under uncertainty and risk we will now focus on the main explicit assumptions. Expected Utility Theory is derived both in the context of risk risk (von Neumann, Morgenstern, 1944) as well as in the context of uncertainty (Savage, 1954, Anscombe, Aumann, 1963), they are stated as properties of a binary relation on the set of probability distributions $L(X)$ in the case of risk and on the set of of acts A in the case of uncertainty. However, in both cases we can distinguish different classes of axioms: three of them are common to both risk and uncertainty and one extra class of axioms is present only in the uncertainty class.

The first class of axioms is the Weak Order axiom applied to the relevant preference relation. This excludes the possibility of indecisiveness (partial orders) or inability to discern the difference between two alternatives (semiorders). According to this axiom all objects of choice (probability distributions or acts) may be ranked according to preference with indifference meaning that the decision maker judges both alternatives equally preferable, and not that she cannot decide. The second class of axioms are continuity axioms. They are usually regarded as technical because in principle they cannot be tested experimentally. However, it is worth noting that there are approaches that relax continuity (see e.g. Kontek, Lewandowski, 2017). The third class of axioms common to risk and uncertainty are the independence axioms. They are crucial for the Expected Utility theories as they force additive separability of the relevant representation and hence impose linearity in probabilities. Since they are so important it's worth to state them here. In the case of risk the independence axiom (also called the substitution axiom, see Kreps, 1988) states the following: for all $P, Q, R \in L(X)$ and $\alpha \in (0, 1)$

$$P \succsim Q \iff \alpha P + (1 - \alpha)R \succsim \alpha Q + (1 - \alpha)R$$

In the case of uncertainty the independence axiom is usually called the sure-thing principle and it states the following: for all $f, g, h, h' \in A$ and $E \subseteq S$

$$(f, A; h, A^c) \succsim (g, A; h, A^c) \iff ((f, A; h', A^c) \succsim (g, A; h', A^c))$$

where $(f, A; h, A^c)$ denotes an act with consequences $f(s)$ for $s \in A$ and $h(s)$ for $s \in A^c$.

Finally, in the context of uncertainty there is one extra class of axioms that force

Michał Lewandowski

separability of tastes (preferences regarding consequences) and beliefs (subjectively perceived likelihood of events). One axiom states that tastes do not depend on events and another states that beliefs do not depend on outcomes used to measure them.

2.3 The EU representation

In the case of choice under risk, von Neumann, Morgenstern (1944) demonstrated how the set of axioms, with Independence being the crucial one, imposed on a preference relation $\succsim \subset L(X) \times L(X)$ is equivalent to the existence of a cardinal utility (unique up to positive linear transformation) function $u : X \to \mathbf{R}$ that represents the preferences in the following sense:

$$P \succsim Q \iff \sum_{x \in \operatorname{supp}(P)} P(x)u(x) \geq \sum_{x \in \operatorname{supp}(Q)} Q(x)u(x)$$

This representation is called the Expected Utility representation.

It is worth noting that the same representation has been already proved before with axioms stated in terms of the so-called quasilinear mean (in economic terms a quasilinear mean corresponds to a Certainty Equivalent of a lottery) instead of a binary relation (see de Finetti, 1931 for the original contribution and Hardy *et al.*, 1934 for the most easily accessible exposition).

Savage (1954) and Anscombe, Aumann (1963) proved an analogous Expected Utility representation for the case of uncertainty. We present here the representation of Savage (1954). They showed that the set of axioms described above imposed on a preference relation $\succsim A \times A$ is equivalent to the existence of a unique non-atomic finitely-additive probability measure μ on S and a non-constant cardinally unique utility function $u : X \to \mathbf{R}$ that represent the preferences in the following sense:

$$f \succsim g \iff \int_S u(f(s))d\mu(s) \geq \int_S u(g(s))d\mu(s)$$

The above representation is called the Subjective Expected Utility representation.

The most important feature of the Expected Utility representation is that the function V_{EU} representing objects of choice (i.e. $P \succsim Q \iff V_{\mathrm{EU}}(P) \geq V_{\mathrm{EU}}(Q)$) is affine in probabilities: for $P, Q \in L(X)$ and $\alpha \in (0, 1)$ the following holds:

$$V_{\mathrm{EU}}(\alpha P + (1 - \alpha)Q) = \alpha V_{\mathrm{EU}}(P) + (1 - \alpha)V_{\mathrm{EU}}(Q)$$

Even though the EU representations allow a general outcome space, it is usually assumed in applications that it is monetary. From now on it will be convenient to denote a typical lottery $P \in L(X)$ by $\mathbf{x} := (x_1, p_1; \dots; x_n, p_n)$, where x_i are monetary outcomes and p_i are the corresponding probabilities (nonnegative, sum up to one). Note that given that the lottery is a probability distribution over X, there is no loss of generality in modeling lotteries as random variables $\mathbf{x}$ instead

of probability distributions P. A random variable is degenerate if the support of its probability distribution contains one element $x \in X$ and will be denoted simply as x or $(x, 1)$. We may also write $(\mathbf{x}, \alpha; \mathbf{y}, 1 - \alpha)$ to denote a compound lottery $\alpha P + (1 - \alpha)Q$, where P, Q are probability distributions of $\mathbf{x}, \mathbf{y}$, respectively. The probability distribution notation is more convenient when discussing the Expected Utility axioms. The random variable notation, on the other hand, is more convenient in further discussion. Similarly, in the context of uncertainty a typical act $f \in A$ will be denoted by $(x_1, E_1; \ldots; x_n, E_n)$, where an event E_i yields outcome x_i. Events E_i are mutually exclusive and exhaustive subsets of S, i.e. they form a partition of S.

2.4   Risk attitudes in the EU representation

The main strength of the EU representation both in its objective and subjective probability form, is its simplicity. Before Expected Utility was introduced, people usually accepted expected value representation. However, it leads to the 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 attitudes. Here are some important examples (For the first two see Rothschild, Stiglitz, 1971, for the third one see Pratt, 1964):

a) $V_{\mathrm{EU}}(.)$ is consistent with the first order stochastic dominance (FOSD) preference iff $u(x)$ is an increasing function. If $\succ_D$ denotes some (partial) order such as FOSD, then we say that $V_{\mathrm{EU}}(.)$ is consistent with this order if and only if $P \succ Q$ whenever $P \succ_D Q$.

b) $V_{\mathrm{EU}}(.)$ exhibits risk aversion (prefers expected value of a lottery to the lottery itself, another saying he dislikes lotteries with zero expected value) or is consistent with the second-order stochastic dominance (SOSD) preference iff $u(x)$ is increasing and strictly concave.

c) $V_{\mathrm{EU}}^*(.)$ is at least as risk averse as $V_{\mathrm{EU}}(.)$ (is willing to pay more to eliminate any risk, namely any lottery with zero expected value) iff its utility function $u^*(.)$ is a concave transformation of $u(.)$.

Risk attitudes are conveniently measured using the two Arrow-Pratt (Pratt, 1964, Arrow, 1964) measures of risk aversion:

a) the absolute risk aversion coefficient: $\mathrm{ARA}(x) = \frac{-u''(x)}{u'(x)}$, for $x \in \mathbf{R}$

b) the relative risk aversion coefficient: $\mathrm{ARA}(x) = \frac{-u''(x)x}{u'(x)}$, for $x \in \mathbf{R}$

Based on these we can define classes of utility functions satisfying certain properties, e.g. (see Pratt, 1964 and Lewandowski, 2013):

a) preferences are shift-invariant iff u belongs to the constant absolute risk aversion (CARA) class, i.e. $u(x) = 1 - e^{-\alpha x}$;

b) preferences are of "wealthier-accept more" type iff u belongs to the decreasing absolute risk aversion (DARA) class;

c) preferences are scale-invariant iff u belongs to the constant relative risk aversion (CRRA) class, i.e. $u(x) = \frac{x^{1-\beta}-1}{1-\beta}$, for $x \geq 0$ if $\beta \neq 1$ and $u(x) = \log(x)$, for $x \geq 0$ as the limiting case when $\beta \rightarrow 1$.

The above is just a few of many similar results. These and many other results allow a very parsimonious and effective modeling of risk attitudes. It has been routinely and successfully applied in a broad range of areas within the economics specialization, not to mention game theory, contract theory, information economics, insurance economics, financial economics, risk measures, general equilibrium models, etc.

The main reason Expected Utility theory is so popular in applications is that it is a normative theory. It is simple and parsimonious and gives strong testable predictions that are consistent with evidence for the most part. It is normative because it sets the standard of rationality. People who violate the Expected Utility axioms either concerning subjective probabilities (de Finetti, 1937) or concerning the expected utility hypothesis (transitivity, independence, see Yaari, 1985) are vulnerable to the so called Dutch books, i.e. sequence of trades that lead to a sure loss of money for the accepting party. Therefore, even if we observe people violating the EU axioms in the laboratory settings, this behavior is likely to vanish or be less persistent in the market settings or other long-term properly incentivized environments allowing learning. The most obvious argument is based on evolution: behavior that leads to a clearly inferior outcomes reduces chances of survival.

2.5 Hidden assumptions concerning the interpretation

Many misunderstandings between the deniers of Expected Utility on one hand and the supporters on the other stems from the differences in defining Expected Utility. We should clearly distinguish between Expected Utility theory, i.e. a mathematical theory based on axioms, with independence being the crucial one and Expected Utility models, i.e. EU theory plus a specific economic interpretation concerning parameters such as X (the outcome space) and S (the state space). This distinction has been advocated by Cox, Sadiraj (2006), Palacios-Huerta, Serrano (2006), Rubinstein (2012), Lewandowski (2014). The standard interpretation that has been associated with Expected Utility by many of its critiques is that of what Rubinstein (2012) calls consequentialism: there exists a single preference relation $\succsim$ on lotteries defined on final (lifetime) wealth levels and preferences $\succsim_W$ on lotteries defined on wealth changes at any given wealth level W are derived from this single preference relation by:

$$\mathbf{x} \succsim_W \mathbf{y} \iff W + \mathbf{x} \succsim W + \mathbf{y}$$

It is important to note that the term consequentialism discussed here has a different (but related) meaning than the doctrine of consequentialism that was discussed at the beginning of section 2.1. The latter term is a very general concept stating that actions are judged solely by their consequences. It does not specify the exact meaning of what consequences mean. Rubinstein's consequentialism assumes that consequences are final wealth positions. It hides yet another implicit assumption, namely that the decision maker starts each new decision problem with all the uncertainty being resolved before. It is reflected in the fact that an initial wealth position W is assumed to be a sure outcome and not a risky prospect. It may well be, however that the initial wealth of a given decision maker is risky. This situation is referred to as background risk (see for instance Gollier, Pratt, 1996).

Consequentialism is surely a valid possibility, but not the only one. In many applications (for example in game theory in general, the discounted expected utility model of a lifetime stream of consumption used to derive the permanent income hypothesis of Friedman, 1957) another interpretation is usually adopted, namely that lotteries are defined on income levels instead of final wealth levels. Cox, Sadiraj (2006) call such model the Expected Utility of income model. It has been explicitly modeled in the framework of Savage as the Reference-dependent Subjective Expected Utility model by Sugden (2003). There are many other possibilities for an interpretation that can be adopted, e.g. Foster, Hart (2009) advocate the so called Expected Utility of gambling wealth model (see also Lewandowski, 2014) in which lottery prizes are integrated with gambling wealth instead of total lifetime wealth. The idea is that people mentally divide their wealth into different budgets (mental accounting), depending on the purpose and time perspective and might decide that they designate some part of a current disposable wealth to gambling. This part is called gambling wealth. Greer, Levine (2006) have a similar suggestion but call the relevant wealth pocket cash money. No matter which interpretation is assumed, the point is to realize that violating consequentialism does not automatically imply violating Expected Utility Theory. However, when applying the model to data, the same axioms but applied within a different economic interpretation may have very different implications. For example under consequentialism transitivity excludes cycles of wealth, whereas under reference dependence with current wealth as reference it does not – in this case it merely excludes cycles of changes of wealth irrespective of the decision maker's current wealth. The above discussion suggests that one should distinguish between the EU principle giving a model the convenient mathematical structure and an economic interpretation, but at the same time remember that EU axioms have different observational implications depending on an interpretation.

It is very important to recognize the influence of the adopted interpretation on the testable predictions of the model. It is necessary to clearly indicate what is the driving force of different testable predictions: either one of the implicit assumptions hidden behind the modeling structure, or one of the mathematical Expected Utility axioms, or finally a given interpretation coupled with the Expected Utility axioms.

As it will turn out a vast part of Expected Utility critique is related to the final wealth consequentialist's interpretation. It means that by adopting a different interpretation and retaining the Expected Utility axioms we could accommodate the evidence that was suggested as evidence against Expected Utility.

3 Expected Utility critique

This section draws from a number of sources. These are: Kahneman, Tversky (1979), Starmer (2000), Machina (1990), Nau (2004), among other. Kahneman, Tversky (1979) started their critique of the standard Expected Utility model with stating the three basic tenets of what they called Expected Utility Theory:

1. the expectation principle: $V(x_1, p_1; \ldots; x_n, p_n) = \sum_{i=1}^{n} p_i u(x_i)$

2. asset integration: $(x_1, p_1; \ldots; x_n, p_n)$ is acceptable at asset position w iff $V(w + x_1, p_1; \ldots; w + x_n, p_n) > u(w)$

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

It is immediately evident that the second (and possibly also third) tenet does not concern the Expected Utility Theory alone but the Expected Utility of wealth model. In some of further works (most notably in Rabin, 2000), it has been implicitly assumed that the asset position w correspond to lifetime wealth level, which leads to the doctrine of consequentialism as defined by Rubinstein (2012).

After stating the main three-tenets of the classical model Kahneman, Tversky (1979) presented evidence against these three basic tenets. This evidence was at the same time meant as a motivation for Prospect Theory - the new model they proposed to replace the standard EU model. The evidence presented by Kahneman, Tversky (1979) and others against the standard Expected Utility model can be divided into three groups:

a) Violations of independence

b) Violations of descriptive and procedural invariance

c) Source dependence (concerns only uncertainty)

Independence is a crucial axiom of Expected Utility Theory ensuring that the indifference curves in the $n-1$-dimensional probability simplex representing lotteries over n outcomes are linear in probabilities, i.e. they are parallel hyperplanes. Violations of independence are thus the primary cause for violations of the Expected Utility principle (see the expectation principle above).

Other axioms of Expected Utility are also commonly violated, but we do not discuss these violations here for the following reasons. The axiom of weak order (completeness and transitivity) is often violated (due to incomparability, imprecise

information, just-noticeable differences, etc.) but these violations are common to contexts other than decision-making under risk and uncertainty such as decisions under certainty, where they are discussed at length. (They differ, however, in the objects they pertain to. In the context of certainty these objects are sure outcomes/consumption bundles, whereas here they are risky/uncertain prospects.)

Violations of the axiom of continuity may well be an important cause for behavior that is not consistent with Expected Utility. However, these violations cannot be tested empirically: continuity requires evaluating arbitrarily small differences between alternatives, but it is questionable/non-testable whether a person's preferences are sensitive toward such changes. Nevertheless, there are models that explain EU paradoxes by relaxing the continuity axiom (Kontek, Lewandowski, 2017).

The second group of violations, namely the violations of descriptive and procedural invariance is by far the largest one (both in terms of published papers as well the number of different kinds of such violations). However, these violations do not concern Expected Utility Theory alone but rather the theory combined with the interpretation of (Rubinstein's) consequentialism.

Finally, source dependence concerns subjective probabilities and are thus confined to the context of uncertainty.

3.1 Violations of the independence axiom

Violations of the independence axiom usually take the form of either of the two effects (Starmer, 2000):

a) Common consequence effect,

b) Common ratio effect.

Common consequence effect can be presented as follows. Suppose there are four compound lotteries:

$$\mathbf{x_1} := (z, p; \mathbf{y}^{**}, 1-p) \quad \mathbf{x_2} := (\mathbf{y}, p; \mathbf{y}^{**}, 1-p)$$
$$\mathbf{x_3} := (x, p; \mathbf{y}^{*}, 1-p) \quad \mathbf{x_4} := (\mathbf{y}, p; \mathbf{y}^{*}, 1-p)$$

where all the lottery supports contain only non-negative outcomes, $p \in (0, 1)$, $x > 0$, the support of $\mathbf{y}$ contains outcomes both greater and less than x, and $\mathbf{y}^{**}$ first order stochastically dominates (FOSD) $\mathbf{y}^{*}$.

People often exhibit the following pattern of preferences: $\mathbf{x_1} \succ \mathbf{x_2}, \mathbf{x_4} \succ \mathbf{x_3}$. This pattern is not consistent with the independence axiom, since it implies the following: $x \succ \mathbf{y} \iff \mathbf{x_1} \succ \mathbf{x_2} \iff \mathbf{x_3} \succ \mathbf{x_4}$.

The most famous version of the common consequence effect is the Allais paradox (Allais, 1953). In the above formulation one needs to make the following substitutions (all consequences are expressed in dollars, M denotes a million): $\mathbf{y}^{**} = 0$, $\mathbf{y}^{*} = 1\mathrm{M}$,

$\mathbf{y} = (5\text{M}, \frac{10}{11}; 0, \frac{1}{11})$, $x = 1\text{M}$ and $p = \frac{11}{100}$. After applying reduction of compound lotteries, one obtains the following two pairs of lotteries:

$$\mathbf{x_1} = (1\text{M}, 0.11; 0, 0.89) \qquad \mathbf{x_2} = (5\text{M}, 0.10; 0, 0.90)$$
$$\mathbf{x_3} = 1\text{M}, \qquad \mathbf{x_4} = (5\text{M}, 0.10; 1\text{M}, 0.89; 0, 0.01)$$

Common ratio effect can be presented as follows: Suppose there are four lotteries:

$$\mathbf{y_1} := (x, p; 0, 1 - p) \qquad \mathbf{y_2} := (y, q; 0, 1 - q)$$
$$\mathbf{y_3} := (x, \alpha p; 0, 1 - \alpha p) \qquad \mathbf{y_4} := (y, \alpha q; 0, 1 - \alpha q)$$

where $1 > p > q > 0$, $0 < x < y$ and $\alpha \in (0, 1)$.

People often exhibit the following pattern of preferences: $\mathbf{y_1} \succ \mathbf{y_2}$, $\mathbf{y_4} \succ \mathbf{y_3}$. This pattern is inconsistent with the independence axiom. Assuming that reduction holds $\mathbf{y_3}$ and $\mathbf{y_4}$ can both be written as compound lotteries: $\mathbf{y_3} = (\mathbf{y_1}, \alpha; 0, 1 - \alpha)$, $\mathbf{y_4} = (\mathbf{y_2}, \alpha; 0, 1 - \alpha)$. Given that, the independence axiom requires: $\mathbf{y_1} \succ \mathbf{y_2} \iff \mathbf{y_3} \succ \mathbf{y_4}$, which is in contradiction with the revealed preferences stated above.

The most common version of the common ratio effect is obtained when (all outcomes are in dollars, K denotes a thousand): $x = 3\text{K}$, $y = 4\text{K}$, $p = 1$, $q = 0.8$, $\alpha = 0.25$. In this case we obtain the following two pairs of lotteries:

$$\mathbf{y_1} := 3\text{K} \qquad \mathbf{y_2} := (4\text{K}, 0.8; 0, 0.2)$$
$$\mathbf{y_3} := (3\text{K}, 0.25; 0, 0.75) \qquad \mathbf{y_4} := (4\text{K}, 0.2; 0, 0.8)$$

Common consequence and common ratio effect exist also if we multiply all the lottery outcomes by -1, obtaining only non-positive outcomes. However, in this case all the preferences are exactly the opposite to what they are above. This phenomenon is called the reflection principle and will be discussed in the subsequent section.

Observe that even though it is possible that common consequence and common ratio lotteries $\mathbf{x_{1-4}}$, $\mathbf{y_{1-4}}$ are all non-degenerate, the most famous examples of these lotteries always involve one lottery that is degenerate: $\mathbf{x_3}$ in the Allais paradox and $\mathbf{y_1}$ in the above version of the common ratio effect. This is not a coincidence. In fact, common ratio and common consequence effects are robust precisely if this is the case. It is closely related to the *certainty effect*. This effect occurs if people overweight consequences that are certain to occur in comparison to those that are not certain. As discussed above violations of independence in the form of either common ratio or common consequence effect are particularly robust if one of the alternatives involved in comparisons is certain. It suggests that the certainty effect may be the main driving force of the violations. For example Conlisk (1989) demonstrates that if lotteries are nudged inside the Marschak-Machina triangle to rule out certainty effect, preference patterns that confirm the Allais and common ratio paradoxes are substantially less prevalent. This finding has also been supported by Harless (1992), Sopher, Gigliotti

(1993), Cohen (1992), Hey, Orme (1994) and Kontek (2015). The phenomenon that is related to the certainty effect is the *possibility effect*. It occurs when people overweight extreme consequences that happen with very small probability.

3.2 Violations of descriptive and procedural invariance

The effects listed below are strongly interrelated with each other.

1. *the isolation effect*: Occurs especially when the decision problem is sequential. In this case, people tend to ignore previous stages when making decisions at subsequent stages. They often disregard common components of the choice alternatives and focus on the differences (Kahneman, Tversky, 1979). However, unlike in the independence axiom, it applies only to obvious differences that are visible to the decision maker right away. Independence requires the same for the differences that are visible only after applying the reduction of compound lotteries. Due to this distinction between the independence axiom and the isolation effect, the latter one may contribute to violations (through the violations of reduction) of the former when gambles are presented in a multi-stage form. Isolation effect also confirms a related phenomenon of reference dependence (see below). It 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.

   a) different representation of probabilities: e.g a choice between the two compound lotteries:

   $$(\mathbf{x}, 0.25; 0, 0.75) \text{ vs. } (\mathbf{y}, 0.25; 0, 0.75)$$

   where $\mathbf{x} = 3\text{K}$ and $\mathbf{y} = (4\text{K}, 0.8; 0, 0.2)$ may be presented as a two stage lottery or as an equivalent (according to reduction) one stage simple lotteries:

   $$(3\text{K}, 0.25; 0, 0.75) \text{ vs. } (4\text{K}, 0.2; 0, 0.75)$$

   May people choose differently in the two choice situations. The modal choice is the LHS lottery in the two stage problem (essentially ignoring the common outcome 0 with probability 0.75 and the RHS lottery in a one-stage problem.
   Another example is that 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.

Michał Lewandowski

---

b) different representation of consequences: e.g. compare the two choice problems: Problem I: you are initially given 1K and then asked to choose between:

$$(+1\text{K}, 0.5; 0, 0.5) \text{ vs. } 500$$

Problem II: you are initially given 2K and then asked to choose between:

$$(-1\text{K}, 0.5; 0, 0.5) \text{ vs. } -500$$

Again, people often choose differently in the two problems (the RHS lottery in Problem 1 and the LHS lottery in Problem 2), even though both the LHS lotteries lead to the same final asset position, and the same is true for the RHS lotteries.

2. *narrow framing*: Narrow framing is the term first introduced by Kahneman, Tversky (1981). It occurs if an agent who is offered a new gamble evaluates that gamble in isolation and does not integrate it with other risks/ does not consider it in a broader decision context.

3. *risky choice framing*: Originally introduced by Kahneman, Tversky (1981), this effect occurs if depending on the description/formulation of different options differing in the level of risk people make different decisions. An example of this effect is the Asian disease paradox: Suppose a community is preparing for the outbreak of an Asian disease which is expected to kill 600 people. You may choose between the following two programs expressed in terms of the number of lives saved:

$$200 \text{ vs. } (600, 1/3; 0, 2/3)$$

Most people choose the sure option. If the same problem is expressed in terms of lives lost:

$$-400 \text{ vs. } (0, 1/3; -600, 2/3)$$

then most people reverse their choice and prefer the risky option.

Another example is that people usually respond differently for identical problems but framed either whether to gamble or whether to insure.

Yet another example of violation of description invariance is that different responses are obtained depending on the description form of a lottery: whether it is a matrix form, decision tree, roulette wheels, written statements, etc.

4. *reference dependence*: Reference dependence occurs if instead of depending on a final asset position choice may also depend on a reference point. The idea of reference dependence (utility defined over nominal wealth changes) was first introduced by Markowitz (1952) in place of the terminal wealth assumption to explain the coexistence of insurance and gambling in the Expected Utility framework (see the classical debate started by Friedman, Savage, 1948).

The EU deniers claim that the terminal wealth assumption was an implicit assumption in most of the EU applications; in fact we do not find support for this claim, there are many areas of application that assume the opposite – utility is defined on wealth changes – see for instance mixed strategy payoffs in game theory or the permanent income hypothesis model of Friedman (1957).

It is worth noting that reference dependence usually refers to having utility defined on the **nominal** changes with respect to the reference wealth level. A related possibility of having utility defined on the **relative** changes with respect to the reference wealth level is present in the classical EU model of wealth: e.g. a preference for maximizing average relative returns relative to some wealth (reserve) level is equivalent to having the logarithmic utility defined on wealth levels – see for instance Foster and Hart (2009) and a series of related papers of Sergiu Hart.

In applications the reference point is usually assumed to be the *status quo* wealth. More generally, it may be defined as the decision maker's recent expectations about outcomes (Kőszegi, Rabin, 2006) or maximin (the max of the minimal outcomes of each lottery under consideration, Schneider, Day, 2016). Good examples of reference dependence are provided in point 1.b) above and 3.

5. *status quo bias*: First introduced by Samuelson, Zeckhauser (1988). It occurs if people prefer things to stay the same as they used to be by doing nothing or by sticking with the decision made previously.

6. *endowment effect*: first demonstrated by Knetsch, Sinden (1984), it occurs when people overvalue a good that they own, regardless of its objective market value. One famous example concerning sure coffee mugs was given by Kahneman *et al.* (1990). Students were randomly assigned to two even groups. Those in one group were given Cornell coffee mugs and those in another not. They could then trade the coffee mugs between them. One could expect that the number of trades should be around half of the number of coffee mugs (those case in which a student who doesn't have a mug values it more than a student who has a coffee mug. It was found that the number of trades was significantly smaller. In fact, the median Willingness-to-pay for a mug was much smaller than the median Willingness-to-accept for it (See 7.a) below).

7. *response mode effects:* described by (Slovic *et al.*, 1982). Different responses depending on whether an experiment was designed to elicit certainty equivalent, gain equivalent or probability equivalent, which under the standard EU model all should yield the same answer.

   a) *WTA-WTP disparity*: People reporting a willingness-to-accept (lowest price for which the DM is willing to sell) much higher than willingness-to-pay (highest price for which the DM is willing to buy) for a given object

Michał Lewandowski

– the ratio of the two prices has been found to be two or more depending on the type of good and experimental procedure. The observed disparities are too large to be reconciled with the standard model. The literature on WTA/WTP disparity is vast starting with Knetsch, Sinden (1984) and summarized partially in Horowitz, McConnell (2002).

b) *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 $P > p$, such that their expected values are close, people often choose the P-bet in a direct choice but assign higher certainty equivalent to the $-bet. An example might be $-bet $=$ $(\$1000, 0.1; 0, 0.9)$ and P-bet $=$ $(\$100, 0.9; 0, 0.1)$. Preference reversal has been studied extensively starting with Lichtenstein, Slovic (1971), Lichtenstein, Slovic (1973) and Grether, Plott (1979).

8. *loss aversion*: means that losses loom larger than gains; people dislike gambles of the form $(x, 0.5; -x, 0.5)$, where $x \neq 0$. Furthermore, if $0 < |x| < |y|$ then people usually express the following preference: $(x, 0.5; -x, 0.5) \succ (y, 0.5; -y, 0.5)$. This phenomenon suggests that there is additional source of aversion to risk than just concavity of the utility function, the latter fact implied by the standard Expected Utility model. Reference dependence with loss aversion is sometimes referred to as gain-loss asymmetry. This property can be used to avoid the implications of local risk-neutrality of the standard EU model (Arrow, 1971) such as the Rabin paradox (see below). Loss aversion may also accommodates the existence of a large gap between Willingness-to-Accept and Willingness-to-Pay for a risky prospect or the preference reversal phenomenon.

9. *The reflection effect*: means that the preference between loss prospects is the mirror image of the preferences between gain prospects. The reflection effect may be combined with the certainty/possibility effect. The two effects together imply the so called four-fold pattern of risk attitudes, i.e. for large probabilities: risk aversion for gains and risk seeking for losses and for small probabilities: risk seeking for gains and risk aversion for losses. Note that this pattern accommodates the phenomenon that people often prefer to gamble and insure at the same time.

10. *Rabin paradox*: Rabin (2000) has shown that in the standard EU model reasonable levels of risk aversion over small stakes imply unrealistically high levels of risk aversion over high stakes. The Rabin's argument is that if an EU agent rejects the equal chance gamble to gain $110 or lose $100 at any initial wealth level, then he will turn down the equal chance gamble to gain an arbitrary sum of money or lose $1000.

The list is non-exhaustive and certainly leaves out many other violations. Besides the effects listed above there is an important class of EU violations that are also robust phenomena. However, we chose to leave them out because we believe that they

correspond to what we might call a mistake - a behavior that is likely to be corrected by the decision maker when the inconsistency is explained to him. Examples include:

a) violations of monotonicity: e.g. in the experiment of Birnbaum, Zimmermann (1998) 70% subjects preferred F over G:

$$F := (96, .85; 90, .05; 12, .10)$$
$$G := (96, .90; 14, .05; 12, .05)$$

This percentage dropped almost to zero when presented as:

$$F := (96, .85; 90, .05; 12, .05; 12, .05)$$
$$G := (96, .85; 96, .05; 14, .05; 12, .05)$$

b) violations of transitivity due to large choice list: when presented with many choices among different risky lotteries, it may well happen that the decision maker makes some intransitive choices not being aware of it.

c) event splitting effects: e.g. Starmer, Sugden (1993) demonstrated that when an event that gives a given outcome is split into two sub-events, there is a tendency for that outcome to carry more weight even though its total probability is unchanged.

3.3   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 (Heath, Tversky, 1991) 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 the 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:

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

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.

4   Prospect Theory

Prospect Theory of Kahneman, Tversky (1979) combines a large number of behavioral phenomena and experimental evidence presented in Section 3 together and proposes a theory which accommodates most of the accumulated evidence against the standard EU model. The theory was originally proposed for risky prospects with up to two distinct outcomes. It is based on two basic tenets:

a) reference dependence,

b) probability distortion.

As discussed in Section 2.5 reference dependence is largely orthogonal to the statement of EU theory. It does not per se violate any of the EU axioms: the EU axioms may be imposed on lotteries defined on changes in wealth, and then the notions may coexist together. However, it is important to note that the EU axioms imposed on lotteries defined on changes in wealth instead of levels of wealth, have different meaning and different normative content than the same axioms imposed on lotteries defined on wealth levels. For example while preferences represented by the EU of wealth model exclude any kind of Dutch books (sequences of trades leading to a sure loss of money) within the considered wealth/budget, preferences represented by the EU of income model may allow Dutch books as demonstrated by Yaari (1985).

While reference dependence is just an alternative economic interpretation of the EU theory, subjective distortion of the objective probability scale concerns the heart of the EU theory, by postulating explicit violation of its main axiom – Independence. The idea was first introduced by Edwards (1954). It has then been applied by Kahneman, Tversky (1979) as one of the building blocks of their theory. It implies that, unlike in EU theory which implies that the indifference curves are linear in probabilities (parallel hyperplanes in a probability simplex), the indifference curves are nonlinear functions of probabilities.

Prospect Theory posits that the choice process involves two phases: editing and evaluation of prospects.

Figure 1: A typical S-shaped value function v with a kink and a typical inverse-S shaped continuous probability weighting function w
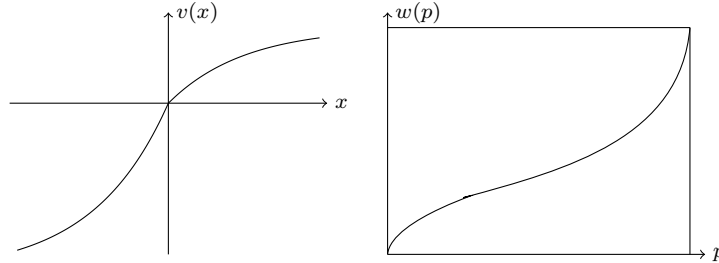
a) *Editing* is meant to serve as a preliminary analysis of a prospect. It specifies rules how to simplify a problem, it involves defining a reference point and hence deciding what is to be regarded as losses and gains, and possibly detecting dominance. This phase is needed to avoid some basic inconsistencies in choice. In the later cumulative version of Prospect Theory, the authors abandoned the idea of the 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 the evaluation phase. However, as Kahneman, Tversky (1979) emphasize, this phase plays an important role in the decision making process and it can account for some oddities in the observed choices.

b) *Evaluation* follows certain rules derived from observed agents' behavior. This part is formalized below.

Reference dependence is modeled by introducing a value function $v : \mathbf{R} \to \mathbf{R}$ which is defined on changes of wealth, for which $v(0) = 0$. Loss aversion is incorporated by requiring that $v(x) < -v(-x)$, for all x. Usually, v has a kink at the origin being a direct indication of the level of loss aversion. The most commonly used form of the value function is the following: given a function $\bar{v}$ defined for $x \geq 0$, with $\bar{v}(0) = 0$, we set $v(x) = \bar{v}(x)$ for gains $x \geq 0$ and $v(x) = -\lambda\bar{v}(-x)$, for all $x < 0$. This value function exhibits loss aversion when $\lambda > 1$. The value function v also exhibits diminishing sensitivity: marginal utility decreases for gains and marginal disutility decreases for losses. A typical value function is depicted in Figure 1. Probability distortion is modeled by the probability weighting function $w : [0,1] \to [0,1]$, which is strictly increasing and satisfies $w(0) = 0$ and $w(1) = 1$. Prospect Theory was designed for binary prospects only. Small probabilities are overweighted ($w(p) > p$, for small p) and moderate and large probabilities are underweighted ($w(p) < p$, for moderate and large p). This implies the inverse-S shaped function. The original function proposed by Kahneman, Tversky (1979) was discontinuous at $p = 0$ and $p = 1$, but it soon was replaced by a continuous function. This function also exhibits diminishing sensitivity,

this time however there are two "starting points" instead of one. Sensitivity towards probabilities diminishes when going away both from $p = 1$ and from $p = 0$ towards the middle.

4.1 Monotonicity problem

One important aspect of a theory is its predictions. It turns out that Prospect Theory with its probability weighting function generates very specific predictions that are related to the issue of monotonicity.

We say that one probability distribution dominates another by First Order Stochastic Dominance if the former can be constructed from the latter exclusively by upward shifts in probability. Alternatively, in the context of uncertainty, a given act dominates another, if for any partition of the state space a consequence of the former act is never worse than that of the latter act for each event in the partition. Intuitively speaking, prospects (or acts) that are dominated are unambiguously worse than the dominating ones. The decision maker's preferences satisfy monotonicity if he never accepts prospect that are dominated. Even though violations of monotonicity occur in experimental settings, they can often be treated as pure mistakes (See discussion in Section 3.2). Many highly respected authors believe that a good theory of choice should not take into account such mistakes. For example Mark Machina 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." (Machina, 1983, p.97).

The implicit assumption of prospect theory is that the decision weight assigned to a specific outcome depends only on the probability of this outcome. It turns out that this assumption generates behavior that intrinsically violates monotonicity wrt to First Order Stochastic Dominance. This will be proved formally in the next Section. Here we provide an example. Suppose that the probability weighting function w underweights probability one half, i.e. $w(0.5) < 0.5$. Consider two prospects $(10 + \epsilon, 0.5; 10, 0.5)$ and $(10, 1)$, where $\epsilon > 0$. It is clear that the former dominates the latter wrt FOSD. Prospect Theory evaluates them, respectively, as $w(0.5)v(10 + \epsilon) + w(0.5)v(10)$ and $v(10)$. It follows by continuity and strict monotonicity of v that we can find $\epsilon > 0$ small enough such that the latter expression is higher than the former, thus violating monotonicity. Similar argument can easily be provided if $w(0.5) > 0.5$. Hence it must be that $w(0.5) = 0.5$ to guarantee monotonicity. However, extending this argument for other than 0.5 probabilities, implies that $w(p) = p$ for all $p \in [0, 1]$. It means that the only probability weighting function that guarantees monotonicity is the one without probability weighting.

5 Rank Dependence-Intuition

This section is based on Diecidue, Wakker (2001). Contrary to the previous section, we will use the framework of uncertainty instead of risk. One reason is that Cumulative

Prospect Theory, unlike original Prospect Theory it replaced, is proposed for both risk and uncertainty. Uncertainty is more general and rank-dependence underlying Cumulative Prospect Theory was first introduced in the context of uncertainty. Therefore it is more natural to discuss it in this context.

As discussed in the previous section the main problem with Prospect Theory is that it does not always satisfy stochastic dominance. In this section it is demonstrated that the central assumptions underlying the Rank Dependent models solve the problem and, at the same time, are very intuitive. The following assumptions are imposed:

A1 *General Weighting Model:* An act $(x_1, E_1; \ldots; x_n, E_n)$ is evaluated according to: $\sum_{i=1}^{n} \pi_i u(x_i)$ where π_i are nonnegative decision weights corresponding to events E_i, such that $\pi_\emptyset = 0$, $\pi_S = 1$ and $u(.)$ is a continuous utility function.

If the decision weights π_i conform to the laws of probability, then the General Weighting Model automatically boils down to Subjective Expected Utility. So we are interested in a more general version, in which decision weights do not necessarily obey the rules of probability. In general, it is well-accepted that monotonicity violations should be excluded by any reasonable theory of choice. Therefore, we impose this as an assumption:

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

Result 1. *Assumptions A1 and A2 imply that weights for any partition of S should sum up to one.*

Proof. By assumption A1, $\pi_S = 1$, so that the claim is true for the one-event partition. We need to prove it for any partition of S. For that we use monotonicity. By A1 $u(x) > u(y) \iff (x, S) \succ (y, S)$ and by A2 $(x, S) \succ (y, S) iff x > y$. Hence u must be increasing. Now consider a partition of S into two events: (E_1, E_2), and suppose that the decision weights for this partition sum up to less than 1 (proof is similar for a sum greater than 1). Then by A1 and continuity of u one can always find $\epsilon > 0$ small enough such that the degenerate act (x, S) is preferred to the act $(x + \epsilon, E_1; x, E_2)$, where $x \in \mathbf{R}$, thus violating monotonicity. Such example can easily be extended for any partition of S. It follows that the decision weights for any partition must always sum up to one. $\square$

The original version of Prospect Theory makes the following assumption which is inherited from the Expected Utility model.

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

Result 2. *Assumptions A1, A2 and A3' imply additivity i.e. for all disjoint events $A, B \subset S$: $\pi_{A \cup B} = \pi_A + \pi_B$.*

Michał Lewandowski

Proof. By A3' we can define for each event E a decision weight $W(E)$. In order to conform to A1 it must be nonnegative, $W(\emptyset) = 0$, $W(S) = 1$ and by A1 and A2 it must sum up to one for any partition of S. Hence $W(E_1 \cup E_2) = W(S) - W(S \setminus (E_1 \cup E_2)) = W(E_1) + W(E_2)$ for any two-element partition (E_1, E_2) for S. It follows that W is a probability measure and the General Weighting Model boils down to SEU. $\square$

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 and requiring no monotonicity violations. Therefore we are interested in relaxing assumption A3'. For that purpose 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: $(x_1, E_1; \ldots; x_n, E_n)$; $x_1 < x_2 < \cdots < x_n$. Define $D_i := E_1 \cup \cdots \cup E_i$, 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. We now relax assumption A3':

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

Result 3. *Assumptions A1, A2 and A3 imply that the decision weight of the maximal outcome is given by a function $W(.)$, depending solely on the event leading to this maximal outcome and satisfying: $W(\emptyset) = 0$, $W(S) = 1$, and $A \subseteq B \Rightarrow W(A) \leq W(B)$.*

Proof. Let a rank-ordered act be defined and denoted as above. By A3 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. By A1 it satisfies $W(\emptyset) = 0$ and $W(S) = 1$. We now prove that it also satisfies monotonicity: $A \subseteq B \Rightarrow W(A) \leq W(B)$. Consider two acts: $(x, A; y, B \cup C)$ and $(x, A \cup B; y, C)$ where $x > y$. By A1, u is strictly increasing, hence: $\pi_{A \cup B} u(x) + \pi_C u(y) \geq \pi_A u(x) + \pi_{B \cup C} u(y)$. 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 up 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 $A \subseteq F$: $W(A) \leq W(F)$. $\square$

The function $W(.)$ defined above is called a capacity. It is in general non-additive, but it satisfies a weaker requirement of monotonicity.

A4 *Solvability:* A capacity $W(.)$ satisfies the following condition:

$$\forall \{A \subset C\} \ \wedge \ \{W(A) \leq p \leq W(C)\} \ \exists B \text{ s.t. } \{W(B) = p\} \ \wedge \ \{A \subset B \subset C\}.$$

Solvability is merely a technical condition that can be regarded as an equivalent of continuity in case of real-valued domains.

Result 4. *Assumptions A1, A2, A3 and A4 imply the so called Choquet expected utility: the rank ordered act $(x_1, E_1; \ldots; x_n, E_n)$ is evaluated according to: $\sum_{i=1}^{n} \pi_i u(x_i)$, where π_i's are defined by:*

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

Proof. We have already defined a decision weight for the highest-outcome event. We need to do the same for other events as well. Consider the following two rank ordered acts:

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

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 $W(E_{i+1} \cup \cdots \cup E_n)$. If we denote the decision weights: π_i^I for the first act and π_i^{II} for the second act, we can write:

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

And it follows directly that:

$$\pi_i = \sum_{j=i}^{n} \pi_j - \sum_{j=i+1}^{n} \pi_j = W(E_i \cup \cdots \cup E_n) - W(E_{i+1} \cup \cdots \cup E_n) \quad \square$$

5.1 Comonotonic independence

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 (i.e. preferences between lotteries or acts are unaffected by substitution of common factors) 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 the 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 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 particularly 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.

6   Cumulative Prospect Theory

This section is based on Tversky, Kahneman (1992). We shall present here the Cumulative Prospect Theory (CPT) 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 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:

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

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

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

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

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

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 decision weights to write:

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

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 nonmonotonic 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):

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

where u is a strictly increasing and continuous reference-dependent value function with $u(0) = 0$.

We now provide a version of CPT for risk and for a general outcome space (allowing both continuous and discrete probability distributions) given by Rieger, Wang (2008).

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

where $F(x) = \int_{-\infty}^{x} dp$ is a cumulative distribution function for outcomes and we require that u is a strictly increasing and continuous reference-dependent value function with $u(0) = 0$ and w_-, w_+ strictly increasing and continuous probability weighting functions such that $w_-(0) = w_+(0) = 0$, and $w_-(1) = w_+(1) = 1$. 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.

Tversky, Kahneman (1992) proposed the following functional form of the value function v and the probability weighting function w:

$$
v(x) = \left\{
\begin{array}{ll}
\frac{x^\alpha}{\alpha}, & x \geq 0, \\
-\lambda \frac{(-x)^\alpha}{\alpha}, & x < 0.
\end{array}
\right.
$$

and the estimated parameter values based on their data is $\alpha = 0.88$ and $\lambda = 2.25$.

$$
w(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}}
$$

and the estimated parameter value is $\gamma = 0.65$.

7 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:

   a) *nonadditive decision weights*

   b) *loss aversion* - by how much losses loom larger than gains

   c) *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 the standard EU model. Two elements of this section are the author's original contributions: the analysis of stochastic dominance in CPT (see Section 7.1) and the discussion on lower and upper subadditivity for risk in Section 7.2).

7.1 Stochastic Dominance in Cumulative Prospect Theory

In this section the lotteries/prospects are represented by cumulative distribution functions F and G. By the classic result under the standard EU model the distribution function F Second Degrees stochastically dominates G if and only if the expected

utility of F is higher than the expected utility of G for all increasing and strictly concave utility functions. In what follows we wish to obtain a similar representation for CPT.

Definition 5. *Given two cumulative distribution functions $F, G : \mathbf{R} \to [0, 1]$, Levy, Wiener (1998) define a Prospect Stochastic Dominance (partial) order $\succsim_P$ on the set of all cumulative distribution functions as:*

$$F \succsim_P G \iff \int_y^x [G(t) - F(t)]dt \geq 0, \ \forall x > 0, \ y < 0. \tag{8}$$

We wish to show under which conditions imposed on the value function u and the rank-dependent probability weighting functions w_+, w_- for gains and losses a CPT model is consistent with the above order. In the proposition below we just show sufficient conditions. Necessary and sufficient conditions may be found in Levy, Wiener (1998).

Proposition 6. *Given two cumulative distribution functions $F, G : \mathbf{R} \to [0, 1]$ the following holds:*

$$F \succsim_P G \Leftarrow V_{CPT}(F) \geq V_{CPT}(G) \tag{9}$$

for any reference-dependent value function u that is convex for losses, concave for gains with $u(0) = 0$ and any rank-dependent probability weighting functions such that w^+ is concave and w^- is convex.

Proof. In the Appendix. □

In what follows we present an example in which $w_-(.)$ and $w_+(.)$ do not satisfy the conditions of the proposition and hence Prospect Stochastic dominance is not retained under the probability transformation. Consider two prospects: $\mathbf{x} = (-2, 1/4; 0, 1/4; 1, 1/2)$ with a probability density function f and $\mathbf{y} = (-1, 1/2; 0, 1/4; 2, 1/4)$ with a probability density function g. As can be seen in the left panel in Figure 2, in the case in which $w_-(.)$ and $w_+(.)$ are identity functions (i.e. there is no probability weighting) F Prospect-Stochastically Dominates G even though neither of the prospects stochastically dominates the other with respect to the First or the Second Degree Stochastic Dominance. In the right panel of Figure 2 the probability weighting functions $w_-(.)$ and $w_+(.)$ are such that $w_-(1/4) = w^+(1/4) = 1/2$ and $w_-(1/2) = w^+(1/2) = 1/2$ (please ignore that it implies that the functions are nonmonotonic). In this figure it is the case that G Prospect- Stochastically Dominates F. Note that the transformed distributions are of the form $\mathbf{x}^* = (-2, 1/2; 0, 1/4; 1, 1/2)$ with probability density function f^* and $\mathbf{y}^* = (-1, 1/2; 0, 1/4; 2, 1/2)$ with probability density function g^*. It is clear that g^* First Order Stochastically dominates $f*$. It is true in general as discussed in the previous sections: if prospect $\mathbf{x}$ First Order Stochastically Dominates prospect $\mathbf{y}$, then prospect $\mathbf{x}$ Prospect-Stochastically Dominates $\mathbf{y}$ as well.

Figure 2: Stochastic Dominance with (right panel) and without (left panel) probability weighting
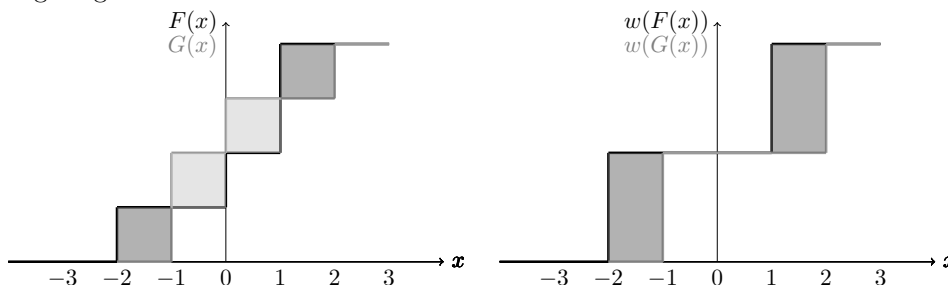
7.2 Probability weighting

This part is based on Tversky, Wakker (1995). 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.

Convex capacity

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, i.e. paying more attention to lower-ranked outcomes, may be defined as follows: if $C \subseteq D$, 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 (i.e. paying more attention to better-ranked outcomes) implies concave capacity, which occurs when the above inequality is reversed.

Lower and upper subadditivity

Definition 7. *A capacity W satisfies subadditivity (SA), if there are events E, E' such that:*

$$W(B) \geq W(A \cup B) - W(A) \quad \text{whenever} \quad W(A \cup B) \leq W(S - E) \tag{10}$$

$$1 - W(S - B) \geq W(A \cup B) - W(A) \quad \text{whenever} \quad W(A) \geq W(E') \tag{11}$$

The condition (10) is called lower SA and the condition (11) 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 $A \succsim B$ if there exists a gain y such that $(y, A) \succsim (y, B)$. Obviously: $A \succsim B$ iff $W(A) \geq W(B)$. We want to show below that conditions (10) and (11) are consistent with the commonly observed choice patterns that we discussed in section 3. In particular we first want to formalize the certainty and the possibility effects.

Definition 8. *We observe certainty effect if:*

$$(x, S - B) \sim (y, A) \Rightarrow (x) \succsim (y, A; x, B), \quad \text{where } 0 < x < y, A \succsim E' \tag{12}$$

where (x) denotes a degenerate act yielding x with certainty, and (x, A) denotes an act yielding x contingent on the occurrence of $A \subset S$ and zero otherwise. More complex acts are denoted in a similar way.

The above statement can easily be derived from the observed choice characteristics in the Allais paradox and from continuity and monotonicity. The other effect leading to nonadditive distortions of subjective probability is *the possibility effect*, i.e. turning impossibility into possibility. This leads to overweighting of small probabilities of extreme events.

Definition 9. *We observe the possibility effect if:*

$$(x) \sim (y, A; x, B) \Rightarrow (y, B; x, S - B) \succsim (y, A \cup B)$$
$$\text{where} \quad 0 < x < y, \ A \cup B \precsim S - E$$

Result 10. *Result on subadditivity: Under the usual requirements, the weighting function W satisfies SA iff (12) and (13) are satisfied.*

Proof. We first show that the upper SA implies the certainty effect. Observe that we can obtain the acts on the RHS of (12) 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 (12) in terms of CPT:

$$u(x)W(S - B) = u(y)W(A) \Longrightarrow$$
$$\Longrightarrow u(x) \geq u(y)W(A) + u(x)(W(A \cup B) - W(A)) \tag{13}$$

Now assume that upper SA holds and multiply both sides of (11) 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 (13) 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.

We now prove that the lower SA implies the possibility effect. Observe that the possibility effect can be derived from the observed choice characteristics and from continuity and monotonicity. Note that we can obtain the acts on the RHS of (13) by changing: B causing x to B causing y in the acts on the LHS. This means that outcome y was impossible 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 (13) in terms of CPT:

$$
\begin{aligned}
u(x) &= u(y)W(A) + u(x)(W(A \cup B) - W(A)) \implies \\
\implies u(y)W(A \cup B) &\leq u(y)W(B) + u(x)(1 - W(B))
\end{aligned}
\tag{14}
$$

Now assume lower SA holds and multiply both sides of (10) by $u(y) - u(x)$ and substitute

$$u(x) = u(y)W(A) + u(x)(W(A \cup B) - W(A))$$

from (14), 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, Wakker (1995) proved also the implication in the other direction. □

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.

Lower and upper subadditivity for risk

Suppose we switch for the moment to the risk situation and imagine we have a probability distortion function $w(.)$ which transforms cumulative probabilities such as the one depicted in Figure 3.

Based on the definition for the case of uncertainty we define lower subadditivity for risk if there exists $\epsilon > 0$ (probability of a lower boundary event) such that $w(p) \geq w(p+q) - w(q)$ whenever $w(p+q) \leq w(1-\epsilon)$. 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: $w'(0) \geq w'(q)$, for any q such that $w(q) \leq w(1-\epsilon)$. This implies that So the function $w(q)$ is at least as steep in $q = 0$ as it is in any $0 < q \leq 1-\epsilon$. The sufficient condition for lower subadditivity is that $w(q)$ is concave for $q \in [0, 1-\epsilon]$. Similarly, upper SA for risk occurs if there exists $\epsilon' > 0$ (probability of an upper boundary event) such that $1 - w(1-p) \geq w(p+q) - w(q)$ whenever $w(q) \geq w(\epsilon')$. 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: $w'(1) \geq w'(q)$ for any q satisfying $w(q) \geq w(\epsilon')$. Thus the function $w(q)$ is at least as steep in $q = 1$ as it is for any $\epsilon' \leq q < 1$. The sufficient condition for upper subadditivity is that $w(q)$ is convex for $[\epsilon', 1]$. Lower and upper subadditivity holds if w is at least as steep at the boundaries (i.e. at 0 and at 1) as in the middle. An example of a probability weighting function satisfying both lower and upper subadditivity is given in Figure 3.

We now focus on a probability weight function that is first concave and then convex (inverse-S shaped). Such a function clearly satisfies lower and upper subadditivity. Assume that $\epsilon' \leq 1 - \epsilon$. In this case, we have a concave region for probabilities in $[0, \epsilon']$, possibly linear region for probabilities $[\epsilon', 1 - \epsilon]$ (if $\epsilon' < 1 - \epsilon$) and a convex region for probabilities $[1 - \epsilon, 1]$. To sum up:

$$w''(p) \begin{cases} \leq 0 & \text{for} \quad p \in [0, \epsilon'] \\ = 0 & \text{for} \quad p \in [\epsilon', 1-\epsilon] \\ \geq 0 & \text{for} \quad p \in [1-\epsilon, 1] \end{cases} \tag{15}$$

For this kind of function it may 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 $[0, 1]$). This can be demonstrated graphically. The right panel of Figure 4 presents the probability weighting functions that satisfy lower and upper subadditivity and yet do not have a fixed point in the interior of the interval $[0, 1]$. In fact they are pretty similar in shape to the entirely

convex and entirely concave functions that do not satisfy subadditivity in the right panel of Figure 4.
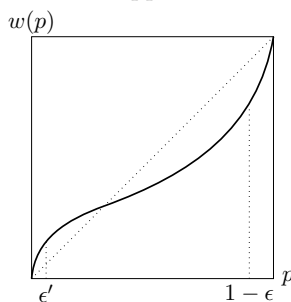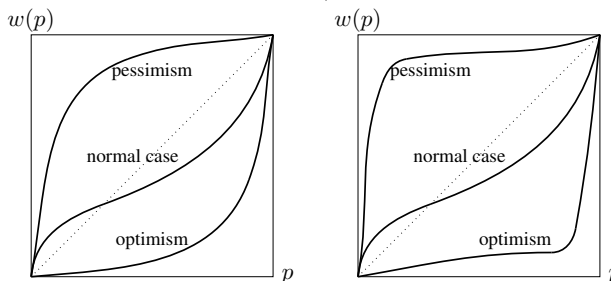
Figure 3: Lower and upper subadditivity in CPT

Figure 4: Pessimism/Optimism in CPT

There are at least two important implications of the above demonstrations:

First, if we want a probability weighting function to exhibit overweighting of both small high-ranked and small low-ranked events, we have to impose additional condition on the weighting function, which will ensure the existence of a fixed point in the interior of the $[0, 1]$ interval. We can simply do it by requiring that $\lim_{p \to 0} w'(p) > 1$ 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: $v_{LSA} := \left[\int_0^{p*} (w(p) - p) dp \right]^+$ ($v_{USA} := \left[\int_{p*}^1 (p - w(p)) dp \right]^+$), where $p*$ is a fixed point of $w(p)$.

Second, the possibility of cases such as described above, i.e. no overweighting of small high-ranked (low-ranked) events and extreme overweighting of small low-ranked (high-ranked) events, suggests a new way of looking at the pessimism and the optimism, i.e. attaching more weight to lower- (higher-, respectively) ranked events. We discussed the notion of pessimism and that of optimism in the paragraph on convex capacities in

this section. Pessimism implies convex whereas optimism implies concave capacity. In the case of risk it is easy to show that pessimism implies concave whereas optimism – convex weighting function. Note that the equivalent of a *convex* capacity in the context of uncertainty is a *concave* probability weighting function in the case of risk.

The above anomalies of the weighting function suggest that a 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 the weighting function is concave almost on the whole domain, but there is no overweighting of lower ranked events and extreme overweighting of higher ranked events. Similarly for a convex probability weighting function. This is presented graphically in Figure 4. The right panel presents an optimistic probability weighting function that is concave on the most part of the interval $[0,1]$, as well as a pessimistic probability weighting function that is convex on the most part of the interval $[0,1]$.

This suggests that we should distinguish a local pessimism/optimism from a global pessimism/optimism. A local pessimism (optimism) implies concavity (convexity) of a weighting function. A 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 v_{LSA} and v_{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.

Comparative subadditivity

After the extensive discussion on the issue of subadditivity, we want to present some results concerning comparative subadditivity and as before, we do it for the case of uncertainty. A transformation $\tau : [0,1] \to [0,1]$ is called SA if it satisfies the same requirements as $W(.)$ does in (10) and (11). 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:

Michał Lewandowski

Result 11. *Result on comparative SA: W_2 is more SA than W_1 iff W_2 is a strictly increasing transform of W_1 and:*

$$W_1(C) = W_1(A \cup B) - W_1(A) \Rightarrow W_2(C) \geq W_2(A \cup B) - W_2(A) \qquad (16)$$

$$\wedge$$

$$1 - W_1(S - C) = W_1(A \cup B) - W_1(A) \Longrightarrow$$
$$\Longrightarrow 1 - W_2(S - C) \geq W_2(A \cup B) - W_2(A) \qquad (17)$$

with boundary condition for lower comparative SA (16): $W_1(A \cup B) \leq W_1(S - E)$ for some E and for upper comparative SA (17): $W_1(A) \geq W_1(E')$ 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. $\qquad \square$

We could also state the corresponding preference conditions for the proposition that $\succsim_2$ is more SA than $\succsim_1$. 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, 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.

7.3 Loss aversion

This part is based on Köbberling, Wakker (2005). 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:

$$u(x) = \begin{cases} U(x) & \text{if } x \geq 0 \\ \lambda U(x) & \text{if } x < 0 \end{cases} \tag{18}$$

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 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. Köbberling, 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.

Kahneman, Tversky (1979) define loss aversion in the following way:

$$\text{for } y > x \geq 0 : \quad u(x) + u(-x) > u(y) + u(-y) \tag{19}$$

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 (19): 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 Köbberling, 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 Kahneman, Tversky (1979, p.279).

8 Cumulative Prospect Theory - Applications and concluding thoughts

This section is based on Camerer (2004). 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, Prescott (1985) showed that under the assumptions of the standard EU model, investors must be extremely risk averse to demand such high a premium, which created a puzzle. Benartzi *et al.* (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 (Shefrin, Statman, 1985). 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 (Shea, 1995). 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 (Bowman *et al.*, 1999).

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 (Thaler, Ziemba, 1988, Hausch, Ziemba, 1995). In case of state lotteries it was observed that large cumulated jackpots attract huge number of people (Clotfelter, Cook, 1993). 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 (Cicchetti, Dubin, 1994). 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 the standard EU model. All the above phenomena can be explained by nonlinear weighting of probabilities, in particular by overweighting of small probabilities of extreme outcomes.

What are the most important situations, in which we can expect to be better off by applying CPT instead of the standard EU model?

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 the standard EU model. 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 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 the standard model. Also, situations which involve constant shifts of reference are likely to generate differences between CPT and the standard EU model predictions, because these shifts change the gain/loss status of outcomes.

Fourth, we should expect larger departures from the standard EU model in situations involving uncertainty rather than risk. 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 the standard EU model. The same argument implies that people such as professional market traders should violate the standard model less often.

The above listing consists of some loose thoughts about the range of applications for CPT. Future research should provide a 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 the standard EU model to CPT provides deeper understanding on the importance of different assumptions underlying the theory. It is however evident that many classic results are not robust to a change from the standard EU model 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

[1] Allais M. (1953). Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'ecole americaine. *Econometrica* 21(4), 503–546.

[2] Anscombe F. J., Aumann R. J. (1963). A definition of subjective probability. *The Annals of Mathematical Statistics* 34(1), 199–205.

[3] Arrow K. J. (1964). The role of securities in the optimal allocation of risk-bearing. *The Review of Economic Studies* 31(2), 91–96.

[4] Arrow K. J. (1965). The theory of risk aversion. [in:] *Aspects of the theory of risk-bearing*, Yrjo Jahnsonin Saatio, Helsinki.

[5] Barberis N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *The Journal of Economic Perspectives* 27(1), 173–195.

[6] Benartzi S., Thaler R. H. (1995). Myopic loss aversion and the equity premium puzzle. *The Quarterly Journal of Economics* 110(1), 73–92.

[7] Birnbaum M. H., Zimmermann J. M. (1998). Buying and selling prices of investments: Configural weight model of interactions predicts violations of joint independence. *Organizational Behavior and Human Decision Processes* 74(2), 145–187.

[8] Bowman D., Minehart D., Rabin M. (1999). Loss aversion in a consumption–savings model. *Journal of Economic Behavior & Organization* 38(2), 155–178.

[9] Camerer C. F. (2004). Prospect theory in the wild: Evidence from the field. [in:] *Advances in behavioral economics*, [ed.:] C. F. Camerer, G. Loewenstein, M. Rabin, Princeton University Press, 148–161.

[10] Cicchetti C. J., Dubin J. A. (1994). A microeconometric analysis of risk aversion and the decision to self-insure. *Journal of Political Economy* 102(1), 169–186.

[11] Clotfelter C. T., Cook P. J. (1993). Notes: The "gambler's fallacy" in lottery play. *Management Science* 39(12), 1521–1525.

[12] Cohen M. (1992). Security level, potential level, expected utility: A three-criteria decision model under risk. *Theory and Decision* 33(2), 101–134.

[13] Conlisk J. (1989). Three variants on the Allais example. *The American Economic Review* 79(3), 392–407.

[14] Cox J. C., Sadiraj V. (2006). Small-and large-stakes risk aversion: Implications of concavity calibration for decision theory. *Games and Economic Behavior* 56(1), 45–60.

[15] De Finetti B. (1931). *Sul concetto di media*. Istituto Italiano degli Attuari.

[16] De Finetti B. (1937). La prévision: ses lois logiques, ses sources subjectives. [in:] *Annales de l'institut Henri Poincaré*, Volume 7, 1–68.

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

[18] Edwards W. (1954). The theory of decision making. *Psychological Bulletin* 51(4), 380–417.

[19] Ellsberg D. (1961). Risk, ambiguity and the savage axioms. *The Quarterly Journal of Economics* 75(4), 643–669.

[20] Foster D. P., Hart S. (2009). An operational measure of riskiness. *Journal of Political Economy* 117(5), 785–814.

[21] Friedman M. (1957). The permanent income hypothesis. [in:] *A theory of the consumption function*, M. Friedman, Princeton University Press, Princeton, 20–37.

[22] Friedman M., Savage L. J. (1948). The utility analysis of choices involving risk. *Journal of Political Economy* 56(4), 279–304.

[23] Gilboa I. (2009). *Theory of decision under uncertainty*. Econometrics Society Monographs, Cambridge University Press, Cambridge.

[24] Gollier C., Pratt J. W. (1996). Risk vulnerability and the tempering effect of background risk. *Econometrica*, 64(5), 1109–1123.

[25] Greer J. M., Levine D. K. (2006). A dual-self model of impulse control. *The American Economic Review* 96(5), 1449–1476.

[26] Grether D. M., Plott C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *The American Economic Review* 69(4), 623–638.

[27] Hardy G., Littlewood J., Pólya G. (1934). *Inequalities.* Cambridge University Press, Cambridge.

[28] Harless D. W. (1992). Predictions about indifference curves inside the unit triangle: A test of variants of expected utility theory. *Journal of Economic Behavior & Organization* 18(3), 391–414.

[29] Hausch D. B., Ziemba W. T. (1995). Efficiency of sports and lottery betting markets. *Handbooks in Operations Research and Management Science* 9, 545–580.

[30] Heath C., Tversky A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty* 4(1), 5–28.

[31] Hey J. D., Orme C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica* 62(6), 1291–1326.

[32] Horowitz J. K., McConnell K. E. (2002). A review of WTA/WTP studies. *Journal of Environmental Economics and Management* 44(3), 426–447.

[33] Kahneman D., Knetsch J. L., Thaler R. H. (1990). Experimental tests of the endowment effect and the coase theorem. *Journal of Political Economy* 98(6), 1325–1348.

[34] Kahneman D., Tversky A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica* 47(2), 263–291.

[35] Kahneman D., Tversky A. (1981). The framing of decisions and the psychology of choice. *Science* 211(4481), 453–458.

[36] Knetsch J. L., Sinden J. A. (1984). Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value. *The Quarterly Journal of Economics*, 99(3), 507–521.

[37] Knight F. H. (1921). *Risk, uncertainty and profit.* Hart Schaffner and Marx, New York.

[38] Köbberling V., Wakker P. P. (2005). An index of loss aversion. *Journal of Economic Theory* 122(1), 119–131.

[39] Kontek K. (2015). Continuous or discontinuous? estimating indifference curves inside the Marschak–Machina triangle using certainty equivalents. *available at:* `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2575526`

[40] Kontek K., Lewandowski M. (2017). Range-dependent utility. *Management Science.* Articles in Advance, doi: 10.1287/mnsc.2017.2744.

[41] Kőszegi B., Rabin M. (2006). A model of reference-dependent preferences. *The Quarterly Journal of Economics* 121(4), 1133–1165.

[42] Kreps D. (1988). *Notes on the Theory of Choice.* Underground classics in economics, Westview Press, Boulder.

[43] Levy H., Wiener Z. (1998). Stochastic dominance and prospect dominance with subjective weighting functions. *Journal of Risk and Uncertainty* 16(2), 147–163.

[44] Lewandowski M. (2013). Risk attitudes, buying and selling price for a lottery and simple strategies. *Central and Eastern European Journal of Economic Modeling and Econometrics* 5, 1–34.

[45] Lewandowski M. (2014). Buying and selling price for risky lotteries and expected utility theory with gambling wealth. *Journal of Risk and Uncertainty* 48(3), 253–283.

[46] Lewandowski M. (2017). Is expected utility an ex-hypothesis? some implications of a reference-dependent Expected Utility model. *available at:* `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3064682`.

[47] Lichtenstein S., Slovic P. (1971). Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology* 89(1), 46.

[48] Lichtenstein S., Slovic P. (1973). Response-induced reversals of preference in gambling: An extended replication in las vegas. *Journal of Experimental Psychology* 101(1), 16.

[49] 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.

[50] Machina M. J. (1990). *Expected utility hypothesis.* [in:] Utility and Probability. The New Palgrave. [ed.:] J. Eatwell, M. Milgate, P. Newman, Palgrave Macmillan, London, 79–95.

[51] Machina M. J., Schmeidler D. (1992). A more robust definition of subjective probability. *Econometrica*, 60(4), 745–780.

[52] Markowitz H. (1952). The utility of wealth. *Journal of Political Economy* 60(2), 151–158.

[53] Mehra R., Prescott E. C. (1985). The equity premium: A puzzle. *Journal of Monetary Economics* 15(2), 145–161.

[54] Nau R. (2004). Phd seminar on choice theory. Lecture notes and readings. *available at:* `https://faculty.fuqua.duke.edu/~rnau/choice/choice04.pdf`.

[55] Palacios-Huerta I., Serrano R. (2006). Rejecting small gambles under expected utility. *Economics Letters* 91(2), 250–259.

[56] Pratt J. W. (1964). Risk aversion in the small and in the large. *Econometrica* 32(1/2), 122–136.

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

[58] Rabin M., Thaler R. (2001). Anomalies. risk aversion. *Journal of Economic Perspectives* 15(1), 219–232.

[59] Rieger M. O., Wang M. (2008). Prospect theory for continuous distributions. *Journal of Risk and Uncertainty* 36(1), 83–102.

[60] Rothschild M., Stiglitz J. E. (1971). Increasing risk II: Its economic consequences. *Journal of Economic theory* 3(1), 66–84.

[61] Rubinstein A. (2012). *Lecture notes in microeconomic theory: the economic agent.* Princeton University Press, Princeton.

[62] Samuelson W., Zeckhauser R. (1988). *Status quo* bias in decision making. *Journal of Risk and Uncertainty* 1(1), 7–59.

[63] Savage L. J. (1954). *The foundations of statistics.* John Wiley & Sons, New York.

[64] Schmeidler D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, 57(3), 571–587.

[65] Schneider M., Day R. (2016). Target-adjusted utility functions and expected-utility paradoxes. *Management Science* Articles in Advance, doi: 10.1287/mnsc.2016.2588.

[66] Shea J. (1995). Myopia, liquidity constraints, and aggregate consumption: a simple test. *Journal of Money, Credit and Banking* 27(3), 798–805.

[67] Shefrin H., Statman M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of Finance* 40(3), 777–790.

[68] Slovic P., Fischhoff B., Lichtenstein S. (1982). Response mode, framing, and information processing effects in risk assessment. [in:] *New directions for methodology of social and behavioral science: Question framing and response consistency* [ed.:] R. Hogarth, Jossey-Base, San Francisco, 21-36.

[69] Sopher B., Gigliotti G. (1993). A test of generalized expected utility theory. *Theory and Decision* 35(1), 75–106.

[70] 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.

[71] Starmer C., Sugden R. (1993). Testing for juxtaposition and event-splitting effects. *Journal of Risk and Uncertainty* 6(3), 235–254.

[72] Sugden R. (2003). Reference-dependent subjective expected utility. *Journal of Economic Theory* 111(2), 172–191.

[73] Thaler R. H., Ziemba W. T. (1988). Anomalies: Parimutuel betting markets: Racetracks and lotteries. *The Journal of Economic Perspectives* 2(2), 161–174.

[74] Tversky A., Kahneman D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5(4), 297–323.

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

[76] von Neumann J., Morgenstern O. (1944). *Theory of games and economic behavior.* Princeton University Press, Princeton.

[77] Wakker P. (2010). Prospect theory: For risk and ambiguity. Cambridge University Press, Cambridge.

[78] Yaari M. E. (1985). On the role of dutch books in the theory of choice under risk. [in:] *Frontiers of Research in Economic Theory: The Nancy L. Schwartz Memorial Lectures 1983–1997*, [ed.:] D. P. Jacobs, E. Kalai, M. I. Kamien, N.L. Schwartz, Cambridge University Press, Cambridge, 33–46.

Appendix

Proof of Proposition 6

STEP 1: We first prove that $V_{\mathrm{CPT}}(F) \geq V_{\mathrm{CPT}}(G)$ for any reference-dependent value function u that is convex for losses, concave for gains with $u(0) = 0$ implies that the

Michał Lewandowski

probability weighting functions w_- and w_+ and the distributions must satisfy the following conditions:

$$\int_0^y [w_+(G(t)) - w_+(F(t))]dt \geq 0, \ \forall y > 0, \ \text{and} \tag{20}$$

$$\int_x^0 [w_-(G(t)) - w_-(F(t))]dt \geq 0, \ \forall x < 0. \tag{21}$$

Integrating by parts all four integrals in (8) the above inequality results in (a, b are the lower and upper bound of the union of supports of F and G; they are allowed to be $-\infty, +\infty$, respectively):

$$[u(x)w_-(F(x))]_a^0 + [u(x)w_+(F(x))]_0^b +$$
$$- \int_a^0 u'(x)w_-(F(x))dx - \int_0^b u'(x)w_+(F(x))dx \geq$$
$$\geq \quad [u(x)w_-(G(x))]_a^0 + [u(x)w_+(G(x))]_0^b +$$
$$- \int_a^0 u'(x)w_-(G(x))dx - \int_0^b u'(x)w_+(G(x))dx$$

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:

$$\left[u'(x)\int_a^x [w_-(G(t)) - w_-(F(t))]dt\right]_a^0 +$$
$$- \int_a^0 u''(x)\int_a^x [w_-(G(t)) - w_-(F(t))]dtdx +$$
$$+ \left[u'(y)\int_0^y [w_+(G(t)) - w_+(F(t))]dt\right]_0^b +$$
$$- \int_0^b u''(y)\int_0^y [w_+(G(t)) - w_+(F(t))]dtdy \geq 0$$

where $x \in [a, 0]$ and $y \in [0, b]$. Rewriting:

$$u'(0) \int_a^0 [w_-(G(t)) - w_-(F(t))]dt +$$

$$- \int_a^0 u''(x) \int_a^x [w_-(G(t)) - w_-(F(t))]dtdx +$$

$$+ u'(b) \int_0^b [w_+(G(t)) - w_+(F(t))]dt +$$

$$- \int_0^b u''(y) \int_0^y [w_+(G(t)) - w_+(F(t))]dtdy \geq 0 \qquad (22)$$

Let us concentrate on the first two terms of the above inequality, decompose the second term and transform as shown below:

$$u'(0) \int_a^0 [w_-(G(t)) - w_-(F(t))]dt +$$

$$- \int_a^0 u''(x) \int_a^x [w_-(G(t)) - w_-(F(t))]dtdx =$$

$$= u'(0) \int_a^0 [w_-(G(t)) - w_-(F(t))]dt +$$

$$- \int_a^0 u''(x) \int_a^0 [w_-(G(t)) - w_-(F(t))]dtdx +$$

$$+ \int_a^0 u''(x) \int_x^0 [w_-(G(t)) - w_-(F(t))]dtdx \qquad (23)$$

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

$$- \int_a^0 u''(x)dx \int_a^0 [w_-(G(t)) - w_-(F(t))]dt =$$

$$= -u'(0) \int_a^0 [w_-(G(t)) - w_-(F(t))]dt + u'(a) \int_a^0 [w_-(G(t)) - w_-(F(t))]dt$$

We can observe that the first element on the RHS of the above equation cancels with the first element on the RHS of equation (23). Going back to the whole inequality

(22), we can write:

$$u'(a) \int_a^0 [w_-(G(t)) - w_-(F(t))]dt +$$

$$+ \int_a^0 u''(x) \int_x^0 [w_-(G(t)) - w_-(F(t))]dt dx +$$

$$+ \quad u'(b) \int_0^b [w_+(G(t)) - w_+(F(t))]dt +$$

$$- \int_0^b u''(y) \int_0^y [w_+(G(t)) - w_+((t))]dt dy \geq 0$$

Because $u''(x) \geq 0$ for $x < 0$ and $u''(y) \leq 0$ for $y > 0$, the above inequality is satisfied whenever $\int_x^0 [w_-(G(t)) - w_-(F(t))]dt \geq 0$ for any $x < 0$ and $\int_0^y [w_+(G(t)) - w_+(F(t))]dt \geq 0$ for any $y > 0$ as claimed.

STEP 2: We need the following two intermediate results: Apart from the usual second order stochastic dominance for risk averse agents, Levy, Wiener, 1998 introduced a similar second order stochastic dominance for risk seeking agents within the EU framework. We list the two results for reference:

a) F dominates G for all risk averse (strictly increasing and concave utility function) EU agents iff $\int_{-\infty}^x [G(t) - F(t)]dt \geq 0$ holds for all x.

b) F dominates G for all risk seeking (strictly increasing and convex utility function) EU agents iff $\int_x^\infty [G(t) - F(t)]dt \geq 0$ holds for all x.

Denote the first kind of domination by SOSD and the second one by SOSD*. Levy, Wiener (1998) have proved that:

1. A monotonic transformation of probability preserves SOSD iff it is concave.

2. A monotonic transformation of probability preserves SOSD* iff it is convex.

Note that in the theorem we require the utility function u to be convex for losses and concave for gains. We can thus use the SOSD domination for gains and SOSD* domination for losses to finish the proof. Suppose that w^+ is concave. Then we know that condition (20) is equivalent to the following condition:

$$\int_0^y [G(t) - F(t)]dt \geq 0 \text{ holds for } y > 0$$

Similarly, suppose that w^- is convex. Then we know that condition (21) is equivalent to the following condition:

$$\int_x^0 [G(t) - F(t)]dt \geq 0 \text{ holds for } x < 0$$

Combining the two conditions we see that:

$$\int_x^y [G(t) - F(t)]dt \geq 0 \ \text{ holds for } x < 0 < y$$

as required.