Prospect Theory for Continuous Distributions Games and Prospects

2 authors, including:

Some of the authors of this publication are also working on these related projects:

Project   What leads to overtrading and under-diversification? Survey evidence from retail investors in an emerging market View project

Prospect Theory for continuous distrbutions

Marc Oliver Rieger   Mei Wang

Prospect Theory for continuous distrbutions

Marc Oliver Rieger and Mei Wang

February 20, 2006

Abstract

Prospect Theory is a widely accepted descriptive framework to model decisions under risk. However, it is limited to situations with finitely many outcomes. Moreover, it is discontinuous, i.e., small changes in a lottery can produce large differences in its utility, contrary to experiments. In this article we present two slight modifications to Prospect Theory which can be theoretically well justified and solve these two problems. We derive in this way a simple model of Prospect Theory for continuous probability distributions that captures all the essential properties of Prospect Theory. Unlike Cumulative Prospect Theory this model has order-independent probability weighting.

1 Introduction

1.1 Prospect Theory

Since Kahneman and Tversky [7] proposed Prospect Theory in 1979, it has accommodated increasing empirical evidence (see, for instance, the review by Starmer [11]), especially when compared with classical Expected Utility Theory, which requires too strict assumptions regarding rationality from the decision makers. Prospect theory adopts the basic framework from Expected Utility Theory, but with additional psychological components based on the observations of the decision making process of real people.

Prospect theory (PT) assumes that decision makers frame outcomes in terms of gains and losses, instead of the final wealth level that is used in expected utility theory (EU). Accordingly, the value function is composed of two part, a concave part in the gain domain and a convex part in the loss domain, capturing the risk-averse tendency for gains and risk-seeking tendency for losses by many decision makers. Another important aspect is that probabilities are weighted by an S-shaped probability weighting function (compare Fig. 1), which is based on the observation that most people tend to overweight small probabilities and underweight large probabilities.

These additional components allow PT to explain violations of some of the properties found in EU (e.g., transitivity and the independence theorem), which have been frequently reported from experiments. However, PT is also criticised for having some undesirable characteristics, especially, the violation of first-order stochastic dominance and continuity. Another limitation of PT is that it can only be applied to discrete outcomes, but applications, e.g. in finance, require a theory for lotteries with non-discrete outcome distributions. To solve these problems, a new theory, Cumulative Prospect Theory (CPT) was proposed in [13], where the *cumulative* probability distributions rather than the probabilities themselves are transformed by the probability weighting function.

Although CPT is often considered as an improvement over PT because it does not violate stochastic dominance and it can be applied to continuous outcomes, the empirical comparisons of these two theories are still inconclusive: some data fit better with PT [3, 15], some data fit better with CPT [4].

Moreover, the studies that aimed at testing the key characteristics of the two theories even seem to suggest frequent contradictions with CPT. Various studies have reported systematic violations of properties of CPT such as ordinal independence, branch independence and event splitting effects [14, 2, 1, 6, 8, 12].

In one recent study using a critical test it has been found that the choices for gambles without a certainty effect are consistent with PT, but not CPT, whereas the choices for gambles with a certainty effect are consistent with both PT and CPT [16].

Given the above evidence, it seems that CPT may be descriptively not as strong as PT. However, PT has some major disadvantages: it does not satisfy first-order stochastic dominance and is also not continuous, and there is no generalization of PT to non-discrete outcome distributions.

We will continue our comparison of PT and CPT at the end of this paper. Before that, we demonstrate in Section 2 that it is possible to extend PT to non-discrete outcome distributions. To this aim, we propose a slight modification of PT, Normalized Prospect Theory (NPT), which enables us to generalize PT to continuous distributions without changing the key features of PT.

In Section 3 we include the concept of the "editing phase" of Kahneman and Tversky [7] into our theory and arrive at a modification of PT that is continuous, and tends to be more compatible with first-order stochastic dominance. We show that this Smooth Prospect Theory (SPT) when normalized leads in the limit to the same generalization for non-discrete distributions as NPT.

In the concluding section we summarize the results of our work and compare the properties of the different theories. In particular, a comparison of NPT with the original experimental data by Kahneman and Tversky demonstrates the descriptive validity of our approach.

1.2   Mathematical preliminaries

In the following sections we need a couple of basic mathematical concepts that we briefly refresh in this introduction:

Let $N \in \mathbb{N}$. We recall that a *probability measure* p on $\mathbb{R}^N$ is a non-negative measure with $\|p\| := \int_{\mathbb{R}^N} dp = 1$. A probability measure is *absolutely continuous* if there exists an integrable function p_a such that we can write

$$\int_{\mathbb{R}^N} f(x) dp(x) = \int_{\mathbb{R}^N} f(x) p_a(x) dx$$

for every continuous function f. A *Dirac* δ_{x_0} is defined by $\int_{\mathbb{R}^N} f(x) \delta_{x_0}(x) = f(x_0)$ for every continuous function f. In particular, we can write any probability distribution p with only finitely many outcomes $x_1, \ldots x_n$ of corresponding probabilities $p_1, \ldots, p_n$ as a sum of Diracs

$$p = \sum_{i=1}^{n} p_i \delta_{x_i}.$$

(We will refer to such sums of Diracs simply as "Dirac measures".) This formulation enables us to handle the two typical situations of discrete lotteries (finitely many outcomes) and continuous outcome distributions (e.g., normally distributed outcomes) simultaneously.

Of course, a probability measure may also be a sum of an absolutely continuous measure and Dirac measures, in other words it might have discrete and continuous parts.

A central tool of this article is the approximation of measures by other measures. To be able to approximate measures, we need to have a notion of convergence. In other words: we want to define when an outcome distribution p is approximated by a sequence of outcome distributions $(p_n)_{n \in \mathbb{N}}$. In order to motivate the mathematical definition, let us consider first the naive approach: we define an approximation by requiring that $p_n(x)$ converges to $p(x)$, i.e. that the probability of every outcome converges. This seemingly natural approach fails for two reasons: first, we are dealing with probability measures for which it is difficult to define $p(x)$ in a reasonable way. (p is not simply a function.) Second, we would exclude that $p_n = \delta_{1/n}$ approximates $p = \delta_0$ which is also not what we want.

There are two better approaches: We could say that p_n converges to p if every expected utility of p_n converges to the expected utility of p. This would imply that, in the limit, no rational investor would see a difference between p and p_n. The other approach is to consider the total probability for an outcome between two arbitrary values x_1 and x_2 (where $x_1 < x_2$). If the total probability for p_n converges (for every choice of x_1, x_2) to the total probability for p, then we could also agree that p_n converges to p.

Both approaches can (under reasonable assumptions) be unified in one definition that coincides with the mathematical concept of weak-$\star$-convergence:

Definition 1.1 (Weak-$\star$-convergence of probability measures). *We say that a sequence $\{p_n\}$ of probability measures on $\mathbb{R}^N$ converges weak-$\star$ to a probability measure p if for all continuous functions f*

$$\int_{\mathbb{R}^N} f(x)dp_n(x) \to \int_{\mathbb{R}^N} f(x)dp(x)$$

holds. We write this as $p_n \overset{\star}{\rightharpoonup} p$. The function f is sometimes called a test function.

To see the correspondence to the two intuitive approaches sketched above, we can consider $f(x)$ as a utility function or (in the second case) as a function which is 1 if $x_1 \leq x \leq x_2$ and zero otherwise.

Finally, we recall the concept of continuity. The word "continuous" unfortunately has two quite different meanings in the English language. Since this may lead to some confusion in this article, we briefly explain the two concepts: First, *continuous* means *non-discrete*. We have already used this notion when talking about measures (or lotteries). As an example, think on a normal distribution in contrast to a lottery with finitely many putcomes. Second, *continuous* means *not discontinuous*. We say that a function F is *continuous* in this sense, if for all sequences x_n converging to x, we have $F(x_n) \to F(x)$. This second type of continuity is an important concept in every model. Roughly spoken, we want a model to be continuously depending on its parameters, since parameters can usually only be measured with a certain amount of precision. This is the case even in a mathematically well-sounded area like physics, and even more so in behavioral decision theory where the precision of experiments is obviously limited. We will discuss the lack of continuity in PT in Section 3. Before that, we turn our attention to the first variant of continuity and generalize PT to continuous (i.e., non-discrete) lotteries which lays the foundation for applications in areas like portfolio optimization.

2   Prospect Theory for continuous probability distribution

We first need to see why there is a difficulty in extending Prospect Theory from discrete to continuous situations. To this end we first consider the naive approach to this problem where we define the PT utility of a continuous distribution p on $\mathbb{R}$ by

$$\widetilde{PT}(p) := \int v(x)w(p(x))\,dx, \tag{1}$$

where v is the value function and w the weighting function.

Every functional that measures preferences has to satisfy a couple of natural conditions which are usually tacitly assumed, one of them being that the utility is invariant under changes of the coordinate system. (The behavior shouldn't change if we "relabel" the monetary units.) However, it is easy to see that the above ansatz (1) does not satisfy this assumption: a rescaling leads to lower values of p and

hence to a different over- or underweighting of the probability distribution. The reason behind this problem is the nonlinearity of the weighting function and hence cannot be circumvented by choosing "appropriate" weighting- or value functions. Let us now try a more sophisticated approach based on an approximation method. The key idea is to approximate the absolutely continuous probability measure p by a sequence of Dirac measures p_n. The usual PT utility for finitely many outcomes can be computed for p_n and we will study the limit $\lim_{n \to \infty} PT(p_n)$. The hope is to find a limit functional that can be used to directly compute $PT(p)$.

There are several problems that we have to overcome here: First, not every possible approximation p_n of p leads to the same limit. This is caused by the highly discontinuous structure of the PT functional. However, we can select a "representative" approximation by formalizing the editing phase as collecting of nearby outcomes. Mathematically spoken, we decompose $\mathbb{R}$ into intervals of equal size $1/n$. On each interval we replace p by a Dirac measure of corresponding weight. More precisely, we define

$$p_{z,n} := \int_{\frac{z}{n}}^{\frac{z+1}{n}} dp, \quad \bar{p}_n := \sum_{z \in \mathbb{Z}} p_{z,n} \delta_{z/n}.$$

The measures $\bar{p}_n$ are still infinite sums of Diracs, but since p is a probability measure, it is easy to see that $\int_{[z/n,(z+1)/n)} dp \to 0$ for $|z| \to \infty$, thus we can neglect all, but finitely many intervals by making an arbitrarily small error. We call the resulting measure p_n. By a small lemma (see [9]), $p_n \overset{\star}{\rightharpoonup} p$.

More general, we can use any decomposition of $\mathbb{R}$ into equally sized intervals $[x_i, x_{i+1})$ where the size $|x_i - x_{i+1}| \to 0$ and the union of these intervals covers all of $\mathbb{R}$ as $n \to \infty$. In fact, to keep the notation simpler, we will use this more general class of approximations from now on.

At this point it is natural to ask why we have to chosen the above approximation. Here we can argue in two ways: first, it is the only *homogenous* approximation that allows for over- and underweighting of the probability in the limit. By "homogenous" we mean that the approximation does not depend in a crucial way on the outcome. This means that the probability weighting of people does not explicitly depend on the outcome involved. (This is, by the way, in contrast to Cumulative Prospect Theory where only the most negative and most positive outcomes are overweighted.)

The second motivation for our approach can be found in [5] where historical data on stocks is used to derive an (approximate) lottery describing their performance which is in turn used to derive subjective PT utilities. The most natural method of forming a lottery is here again to integrate the (discrete) events into a histogram that corresponds to a probability distribution. This approach is based on the understanding of the convergence properties of PT utilities in the limit of infinitely many data points approximating an underlying probability distribution. The results presented here provide this basis.

We can now study the limit of the resulting PT utilities and we will show the following (at first glance a little surprising) result:

Theorem 2.1. *Let p be a continuous probability distribution on $\mathbb{R}$ with expected utility*

$$EU(p) := \int v(x)p(x)\,dx \neq 0$$

and let p_n be defined as above. Assume $v \in C^1(\mathbb{R})$. Moreover, assume that for the weighting function $w \colon [0,1] \to [0,1]$ there exists some $\alpha \in (0,1]$ and some $C > 0$ such that

$$w(\varepsilon) = C\varepsilon^{\alpha} + O(\varepsilon^2) \tag{2}$$

for $\varepsilon \to 0$. Then the PT utility

$$PT(p_n) := \sum_z w(p_{n,z})v(z/n)$$

converges to

$$\lim_{n\to\infty} PT(p_n) = \begin{cases} \infty, & \text{if } \alpha < 1, \\ C \cdot EU(p), & \text{if } \alpha = 1, \end{cases}$$

Proof. Since p is absolute continuous, $p_{i,n} := \int_{x_i}^{x_{i+1}} p(x)\,dx$ converges to zero as $n \to \infty$. Hence we can use the expression (2) to prove

$$
\begin{aligned}
\lim_{n\to\infty} PT(p_n) &= \lim_{n\to\infty} \sum_i w(p_{i,n})v(x_i) \\
&= \lim_{n\to\infty} \sum_i C(p_{i,n})^{\alpha} v(x_i). \tag{3}
\end{aligned}
$$

In the case $\alpha = 1$, we derive from this

$$\lim_{n\to\infty} PT(p_n) = \lim_{n\to\infty} \sum_i C\left(\int_{x_i}^{x_{i+1}} v(\xi)\,dp + \int_{x_i}^{x_{i+1}} (v(x_i) - v(\xi))p(x)\,dx \right).$$

Since $|v(x_i) - v(\xi)| \le |v'(x_i)||x_i - \xi| + O(|x_i - \xi|^2) \to 0$ as $n \to \infty$, for every converging sequence of x_i, we get

$$
\begin{aligned}
\lim_{n\to\infty} PT(p_n) &= C \lim_{n\to\infty} \sum_i \int_{x_i}^{x_{i+1}} v(\xi)p(x)\,dx \\
&= C \cdot EU(p).
\end{aligned}
$$

We now consider the case $\alpha < 1$. From estimate (3) we obtain

$$\lim_{n\to\infty} PT(p_n) = \lim_{n\to\infty} \sum_i C\frac{p_{i,n}}{(p_{i,n})^{1-\alpha}}v(x_i).$$

We estimate $(p_{i,n})^{1-\alpha} \le (s_n)^{1-\alpha}$ with $s_n := \sup_i p_{i,n}$. Since $p_{i,n} \to 0$ as $n \to \infty$, we have $s_n \to 0$ as $n \to \infty$. Therefore we arrive at

$$
\begin{aligned}
\lim_{n\to\infty} PT(p_n) &\ge C \lim_{n\to\infty} s_n^{\alpha-1} \sum_i p_{i,n}v(x_i) \\
&= C \lim_{n\to\infty} s_n^{\alpha-1} EU(p),
\end{aligned}
$$

which is infinite, since $EU(p) \neq 0$. $\qquad\square$

Remark 2.2. *As we have seen, the condition $EU(p) \neq 0$ is only used in the case $\alpha < 1$. Since the expected utility is only meaningful up to an affine transformation, this can be assumed without loss of generality.*

Remark 2.3. *The classical case studied by Kahneman and Tversky [13] is the case of $\alpha < 1$. A weighting function with $\alpha = 1$ has been suggested in [10]. We could also consider the case $\alpha > 1$. Here the limit of $PT(p_n)$ as $n \to \infty$ is zero. The proof is similar to the above cases, but the result is less interesting, since this case corresponds to an* underweighting *of low probabilities which is experimentally not observed.*

Theorem 2.1 highlights the difficulty of probability weighting in Prospect Theory; in the approximation process, the single probabilities become smaller and smaller, hence (if $\alpha < 1$) the overweighting becomes stronger and stronger and finally leads to an infinite utility. In the case $\alpha = 1$, however, the relative difference between the overweighing becomes smaller and smaller as the single probabilities become small, hence in the limit the overweighting does not play a role anymore and we arrive simply at a variant of the expected utility.

Both cases are of course useless: we need a finite limit functional that still reflects the basic property of probability weighting. We therefore need to renormalize the PT functional in order to avoid the infinite overweighting. This is reasonable also from a modeling point of view: a person might overweight certain unlikely events, however, it is somehow strange to think that a person will end up with a perceived *total probability* larger than one. We might rather expect that an overweighting of *one* event will always cause the underweighting of *another* event. Classical Prospect Theory negates this idea, whereas Cumulative Prospect Theory always has a total weighted probability of one. This difference does not matter much for lotteries with only a couple of outcomes, but it becomes crucial in our approximation where the number of outcomes tends to infinity.

For these reasons we suggest the following normalization of the classical PT functional that sets the sum of the weighted probabilities of PT to one:

$$NPT(p) := \frac{\sum_{i=1}^{n} w(p_i)v(x_i)}{\sum_{i=1}^{n} w(p_i)}.$$

We will study some of the properties of this functional later. We will also show that this normalized theory gives comparable results to the usual PT when applied to data from classical lottery experiments. For the moment we want to demonstrate that this new approach enables us to derive, using the above convergence method, a variant of Prospect Theory for continuous probability distributions.

Theorem 2.4. *Let p be a probability distribution on $\mathbb{R}$ with exponential decay at infinity and let p_n be defined as above. Assume that $v \in C^1(\mathbb{R})$ has at most polynomial growth and that for the weighting function $w \colon [0,1] \to [0,1]$ there exists some $\alpha \in (0,1)$ and some $C > 0$ such that $w(\varepsilon) = C\varepsilon^\alpha + O(\varepsilon^2)$ for $\varepsilon \to 0$.*

Then the normalized PT utility

$$NPT(p_n) := \frac{\sum_z w(p_{n,z})v(z/n)}{\sum_z w(p_{n,z})}$$

converges to

$$\lim_{n\to\infty} NPT(p_n) = PT(p) := \frac{\int v(x)p(x)^\alpha \, dx}{\int p(x)^\alpha}.$$

(We can call the limit PT(p), since Prospect Theory has previously been only defined for discrete lotteries, but p is non-discrete, in fact absolutely continuous.)

Proof. The proof follows the same general ideas as the proof of Theorem 2.1. We denote $h_n := |x_i - x_{i+1}|$. (Remember that, since the decomposition is assumed to be homogenous, h_n does not depend on i.)

$$
\begin{aligned}
\lim_{n\to\infty} NPT(p_n) &= \lim_{n\to\infty} \frac{\sum_i w(p_{i,n})v(x_i)}{\sum_{i=1}^n w(p_{i,n})} \\
&= \lim_{n\to\infty} \frac{\sum_i C(p_{i,n})^\alpha v(x_i)}{\sum_i C(p_{i,n})^\alpha} \\
&= \lim_{n\to\infty} \frac{h_n^\alpha \sum_i \left(\int_{x_i}^x x_{i+1}p(x)\,dx\right)^\alpha v(x_i)}{h_n^\alpha \sum_i \left(\int_{x_i}^x x_{i+1}p(x)\,dx\right)^\alpha}.
\end{aligned}
$$

We see that we can now rescale the enumerator by h_n^α (contrary to Theorem 2.1), since we have another integral expression in the denominator. In the next step, we transform the integrals into averages of which we can finally take the limit $n \to \infty$ (since p is continuous):

$$
\begin{aligned}
\lim_{n\to\infty} NPT(p_n) &= \lim_{n\to\infty} \frac{\sum_i \left(\fint_{x_i}^{x_{i+1}} p(x)\,dx\right)^\alpha v(x_i)}{\sum_i \left(\fint_{x_i}^{x_{i+1}} p(x)\,dx\right)^\alpha} \\
&= \frac{\int v(x)p(x)^\alpha \, dx}{\int p(x)^\alpha}.
\end{aligned}
$$

This concludes the proof of Theorem 2.4.   □

Remark 2.5. *The assumption that p has exponential decay at infinity and that v has at most polynomial growth is needed in order to ensure that $\int p(x)^\alpha \, dx$ and $\int p(x)^\alpha v(x)\,dx$ both have finite values. (The assumption can be weakened, e.g. to $\int p(x)^\alpha \, dx < +\infty$ if v is bounded.) This problem is closely related to the St.-Petersburg paradox that occurs in CPT and is interesting on its own. The curious reader may compare this with the results in [10].*
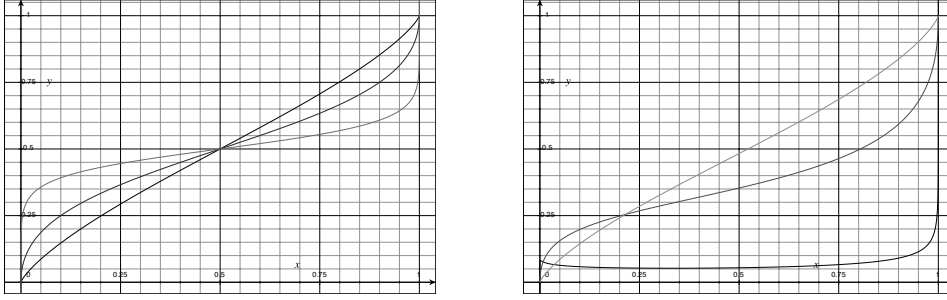
Figure 1: Effective weighting function for lotteries with *two* outcomes of NPT compared to the classical weighting function by Kahneman and Tversky [13]. Both functions are displayed for $\alpha = 0.2$, 0.5, 0.8.

It is surprising that this approach has not been used before. A reason for that might be that at first glance there seems to be no rationale why the sum of the overweighted probabilities should be one. The standard example is here a lottery where two events with the small probability p show a stronger influence on a decision than one event with probability $2p$. However, a closer look reveals that such experiments do not contradict our approach since it is again only about the *relative* overweighting of the two events of the former lottery comparing to the single event of the latter. The remaining outcomes are not taken into account and we expect a relative *underweighting* of them.

Another potential problem with the approach might be that the resulting continuous functional seems to contradict a core ingredients on PT, namely the S-shape of the weighting function. In fact, the probability density $p(x)$ is only weighted as $p(x)^\alpha$, so we have a strictly concave "probability weighting function" and no S-shape. However, this is not really the case: the normalization ensures that large proabilities are still underweighted. We can see this effect directly in the discrete case of NPT:

Example 2.6. *Let us chose the weighting function w in NPT as $w(p) := p^\alpha$ (which is the natural case as it corresponds to the weighting function we get in NPT for continuous distributions). We consider the case of* two *lotteries. Then the NPT subjective utility corresponds to the PT subjective utility for the S-shaped weighting function*

$$w(p) := \frac{p^\alpha}{(p^\alpha + (1-p)^\alpha)}.$$

The proof of this is straightforward, hence we omit it here. We sketch the graph of w in Fig. 1 and compare it with the standard weighting function by Kahneman and Tversky [13]. The function w is, unlike the original suggestion of Kahneman and Tversky, monotone increasing for all $\alpha > 0$, compare [10].

At this point one might wonder why the (symmetric) shape of Fig. 1 does not exclude an explanation of the classical Allais paradox (compare Table 1). It is one of the features of PT that it is able to explain the decision process leading to

Lottery A	Probability	33%	66%	1%
NPT=227.4	Outcome	2500 €	2400 €	0 €
Lottery B	Probability	100%		
NPT=232.4	Outcome	2400 €		
Lottery C	Probability	33%	67%	
NPT=90.5	Outcome	2500 €	0 €	
Lottery D	Probability	34%	66%	
NPT=89.7	Outcome	2400 €	0 €	

Table 1: In the Allais paradox, people typically prefer Lottery B over Lottery A and Lottery C over Lottery D, thus violating the independence axiom. This is predicted correctly by NPT.

the paradox, based on the violation of the independence axiom (the "independence of irrelevant alternatives"). It is easy to see that PT with a symmetric weighting function (i.e. $w(p) = 1 - w(1 - p)$) does not allow an explanation of the Allais paradox. However, the "effective" weighting function of Fig. 1 is only valid for two outcomes. In order to formulate the Allais paradox, we need at least three different outcomes. In this case it is not possible to rewrite NPT as a variant of PT, thus there is no "effective" weighting function. In fact, we can show that the Allais paradox can still be explained with NPT: assume that $w(p) = p^\alpha$ and $v(x) = x^\beta$ with $\alpha = \beta = 0.7$, then a short computation gives the NPT-values shown in Table 1, thus NPT predicts, like PT, the experimentally observed preferences.
In Section 4 we will test the classical experimental data by Kahneman and Tversky [13] with NPT and the simple weighting function $w(p) = p^\alpha$.

3   Continuity in Prospect Theory

Prospect Theory lacks some properties which would be quite natural to assume. In particular it violates stochastic dominance and is discontinuous. In the classical framework of PT, there is therefore a process called "editing phase" which "filters" such problems. The person is assumed to process first the presented lotteries, in particular by collecting outcomes with identical (or nearly identical) values to a single event. As an example for this editing process consider the following two lotteries:

outcome	0€	9.99€	10€
probability	0.8	0.1	0.1

outcome	0€	10€
probability	0.8	0.2

Without an editing phase, PT could value the first higher than the second, due to the strong overweighting of the low probability 0.1. In the editing phase, however, the first lottery would be converted into a lottery similar to the second one, by simply collecting the very similar payoffs of 9.99€ and 10€.

In this example we see two effects of the editing phase: on the one hand, it avoids stochastic dominance violation. (The second lottery is "obviously better", i.e. it is stochastically dominant.) On the other hand, it avoids a discontinuity of the theory: a sequence p_n of lotteries of the first type that converges to the second lottery p does in PT (without editing phase) not satisfy the continuity condition $PT(p_n) \to PT(p)$.

In this section we present a modification of PT that incorporates this basic idea of the editing phase. We will demonstrate that it is compatible with the normalized PT introduced in the last section in that (after renormalization) its limit functional for non-discrete probability distributions coincides with the limit functional of NPT.

At the end of this section we will then compare certain properties of this modified version, called Smooth Prospect Theory (SPT), with PT and CPT to discuss its usefulness.

Our idea is to model the editing phase by making our look on the payoffs a little bit "fuzzy". We introduce a parameter $\varepsilon > 0$ and assume that outcomes that differ only by ε or less are more and more considered to be the same. In a first step we therefore transform $p := \sum_{i=1}^{n} p_i \delta_{x_i}$ into an absolute continuous probability distribution p_ε by the formula

$$p_\varepsilon(x) := \sum_{i=1}^{n} p_i \chi_{[x_i - \varepsilon, x_i + \varepsilon]}(x),$$

whenever $\chi_{[a,b]}(x) = 1$ where $a \le x \le b$ and $\chi_{[a,b]}(x) = 0$ elsewhere. ($\chi_{[a,b]}$ is called the "characteristic function" of the interval $[a,b]$.) In Fig. 2 we give an illustration for such a transformation, based on a lottery similar to the one from our initial example. (The idea is closely related to the concept of "kernel estimates" and to "mollifiers" in analysis.)
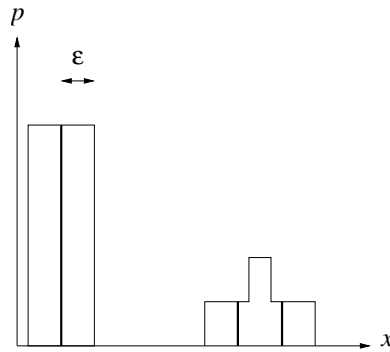

Figure 2: A lottery before (thick lines) and after (thin lines) the "editing phase" as described by the SPT model.

Given this transformed probability, we now need to define its subjective utility in a way that we recover the classical PT when $\varepsilon \to 0$. To this aim, we need to define

the *Smooth Prospect Theory (SPT)* of the lottery p by

$$SPT_\varepsilon(p) := \int \frac{1}{2\varepsilon} w\left(\sum_{i=1}^{n} p_i \chi_{[x_i-\varepsilon, x_i+\varepsilon]}(x)\right) v(x)\, dx \tag{4}$$

(In the following, we will occasionally omit the index ε. Then ε is an arbitrary fixed positive number.) We first prove that the above definition makes sense:

Proposition 3.1. *Let $SPT_\varepsilon(p)$ be given by (4). Then $\lim_{\varepsilon\to 0} SPT_\varepsilon(p) = PT(p)$. Moreover, SPT_ε is invariant under affine rescaling.*

Proof. If $\varepsilon > 0$ is smaller than $\min_{i,j} |x_i - x_j|$, then (4) simplifies to

$$
\begin{aligned}
SPT_\varepsilon(p) &= \int \frac{1}{2\varepsilon} \sum_{i=1}^{n} w(p_i)\chi_{[x_i-\varepsilon, x_i+\varepsilon]}(x)v(x) \\
&= \sum_{i=1}^{n} w(p_i) \int_{x_i-\varepsilon}^{x_i+\varepsilon} v(x)\, dx.
\end{aligned}
$$

Since $v \in C^1$, we obtain $\lim_{\varepsilon\to 0} SPT_\varepsilon(p) = \sum_{i=1}^{n} w(p_i)v(x_i)$.

A straightforward computation finally shows that SPT is invariant under affine rescaling of the monetary units.   □

The main purpose of incorporating the editing phase into the mathematical formalism was to avoid the discontinuity of the original theory. The next theorem shows that SPT is in fact continuous:

Theorem 3.2. *Let p_n and p be Dirac measures with $p_n \stackrel{\star}{\rightharpoonup} p$, then $SPT(p_n) \to SPT(p)$, i.e., Smooth Prospect Theory is continuous.*

Proof. First, we notice that (4) can be rewritten as follows:

$$
\begin{aligned}
SPT_\varepsilon(p) &= \int \frac{1}{2\varepsilon} w\left(\sum_{i=1}^{n} p_i \chi_{[x_i-\varepsilon, x_i+\varepsilon]}(x)\right) v(x)\, dx \\
&= \int \frac{1}{2\varepsilon} w\left(\int_{x-\varepsilon}^{x+\varepsilon} dp\right) v(x)\, dx. \tag{5}
\end{aligned}
$$

Using the definition of weak-$\star$ convergence and the test function $\chi_{[x-\varepsilon, x+\varepsilon]}$ we obtain for $p_n \stackrel{\star}{\rightharpoonup} p$ that

$$\int_{x-\varepsilon}^{x+\varepsilon} dp_n \to \int_{x-\varepsilon}^{x+\varepsilon} dp.$$

Hence $SPT(p_n) \to SPT(p)$.   □

Remark 3.3. *Although SPT is continuous, it can still violate the stochastic dominance principle if $\varepsilon > 0$ is chosen small enough.*

In fact, one can show that for every fixed $\varepsilon > 0$, there can be lotteries violating the stochastic dominance principle. Hence the "collecting" of similar outcomes by itself is not a sufficient explanation for the avoidance of dominated lotteries. (The proof of this is relatively easy, one just needs to construct a lottery with outcomes being apart at least ε and so low probabilities that the overweighting still leads to a stochastic dominance violation similar to the initial example.) However, the number of stochastic dominance violations decreases, the larger the ε, the more.

We can now study the same limit for SPT that we have studied for normalized PT in the previous section. Again, we first have to normalize the total subjective probability to one. We write

$$SNPT(p) := \frac{\int w\left(\sum_{i=1}^{n} p_i \chi_{[x_i-\varepsilon, x_i+\varepsilon]}(x)\right) v(x)\, dx}{\int w\left(\sum_{i=1}^{n} p_i \chi_{[x_i-\varepsilon, x_i+\varepsilon]}(x)\right) dx}.$$

With this definition we obtain the following result:

Theorem 3.4. *Let p be a probability distribution on $\mathbb{R}$ with at least exponential decay at infinity and let p_n be probability measures with $p_n \overset{\star}{\rightharpoonup} p$. Assume that $v \in C^1(\mathbb{R})$ has at most polynomial growth and that for the weighting function $w\colon [0,1] \to [0,1]$ there exists some $\alpha \in (0,1)$ and some $C > 0$ such that $w(\delta) = C\delta^\alpha + O(\delta^2)$ for $\delta \to 0$. Moreover, let $\varepsilon_n \to 0$ as $n \to \infty$. Then the SNPT utility of p_n converges to*

$$\lim_{n\to\infty} SNPT_{\varepsilon_n}(p_n) = PT(p) = \frac{\int v(x) p(x)^\alpha\, dx}{\int p(x)^\alpha}.$$

Proof. We use the representation (5) for the definition of $SNPT$ and multiply denominator and enumerator each by $\varepsilon_n^{-\alpha}$ to obtain

$$SNPT(p_n) = \frac{\int \varepsilon_n^{-\alpha} w\left(\int_{x-\varepsilon_n}^{x+\varepsilon_n} dp_n\right) v(x)\, dx}{\int \varepsilon_n^{-\alpha} w\left(\int_{x-\varepsilon_n}^{x+\varepsilon_n} dp_n\right) dx}.$$

Since $\varepsilon_n \to 0$ as $n \to \infty$, and the limit measure p is absolutely continuous, the integrals $\int_{x-\varepsilon_n}^{x+\varepsilon_n} dp_n$ converge to zero (uniformly in x). Using the growth of w near zero, we obtain

$$\lim_{n\to\infty} SNPT(p_n) = \lim_{n\to\infty} \frac{\int \left(\varepsilon_n^{-1} \int_{x-\varepsilon_n}^{x+\varepsilon_n} dp_n\right)^\alpha v(x)\, dx}{\int \left(\varepsilon_n^{-1} \int_{x-\varepsilon_n}^{x+\varepsilon_n} dp_n\right)^\alpha dx}.$$

Arguing like in the proof of Theorem 2.1 and using the weak-$\star$ convergence of p_n, this proves the convergence to $PT(p)$. $\qquad\square$

Remark 3.5. *It is easy to verify that corresponding variants of Proposition 3.1, Theorem 3.2 and Remark 3.3 hold for SNPT as well.*

In the following section we discuss the properties of the various variants of Prospect Theory and compare them with experimental findings.

Properties	PT	NPT	SNPT	CPT	Empirical Evidence
Violation of transitivity	no	no	no	no	yes

Violation of internality
  Example: when v(x)=x and w(.25)=.4,
  the CE of (.25,100;.25,90;.25,.80;.25,0)
  is 108, greater than any of the outcomes.
	yes	no	no	no	no

Violation of continuity	yes	yes	no	no	
Violation of stochastic dominance	yes	yes	yes	no	
Violation of independence axiom	yes	yes	yes	yes	yes
Inverse S-shaped weighting function which implies lower- and upper-subadditivity	yes	yes	yes	yes	yes
Extreme outcomes receive more weights (rank-dependence)	no	no	no	yes	yes
Distinctive outcomes receive more weights (support theory)	yes	yes	yes	no	yes
Among distinctive outcomes, closer outcomes get less weights	no	no	yes	no	
Can be applied for continuous distribution	no	yes	yes	yes	

Table 2: Comparisons of Prospect Theory, Normalized Prospect Theory, Smooth Normalized Prospect Theory, and Cumulative Prospect Theory,

4   Descriptive validity

Prospect Theory and Cumulative Prospect Theory have been accepted as the most competitive alternative theories of Expect Utility Theory to describe decision under risks. Although CPT is often considered to have mathematically more elegant properties, empirical evidence sometimes suggest that the original PT may capture certain psychological processes that cannot be predicted by CPT. Unfortunately, the application of PT is only limited to discrete distributions, which motivated us to propose a slight modification, Normalized PT which can be extended to continuous setting. The lack of continuity in the classical PT led us to another modification that tries to include the editing phase or more specifically the aggregation of close outcomes. The result is Smooth Prospect Theory that can again be normalized.

We compare several mathematical properties of PT, NPT, SNPT and CPT to the empirical evidence, see Table 2, compare also [16]. It seems that NPT preserves most of the properties of PT, except for the violation of internality, which is (viewed from the experimental evidence) an undesirable property of PT. SNPT outperforms PT and NPT in that it does not violate continuity and predicts that closer outcomes are less overweighted than very distinct outcomes, which is psychologically very plausible. Compared to other PT theories, the rank-dependent property of CPT is

Gains	PT	NPT	Losses	PT	NPT
parameter for weighting function	0.68	0.59	parameter for weighting function	0.71	0.94
parameter for value function	0.83	0.66	parameter for value function	0.64	0.77
goodness of fit	97.2%	97.6%	goodness of fit	98.3%	98.3%

Table 3: Comparison of the average of the fitted parameters and the goodness of fit between PT and NPT for gains (left) and losses (right).

more consistent with empirical evidence, however, it fails to predict that distinct outcomes receive more weights than aggregated outcomes, which has been found in experiments. In some applications, in particular in behavioral finance, an advantage of all variants of classical PT (thus also NPT and SNPT) is the reduced computation load as compared to CPT, since an ordering of the probabilities by their outcomes is not necessary.

After this qualitative comparison, we present the results of a quantitative analysis of the NPT model. We reevaluate the classical data set by Tversky and Kahneman [13] and compare PT (with the value- and weighting functions as originally suggested) with NPT where we choose the value function as in [13] and take the natural weighting function $w(p) := p^\alpha$ which we obtain in the limit of continuous distributions. Both models have therefore two fitting parameters (in gains and in losses each), however, the NPT model has no flexibility in choosing the weighting function (besides adjusting the parameter α), whereas in PT the type of the function has to some extent been chosen in order to accommodate the experimental data. We use a nonlinear regression with R^2 as measurement of fit for each individual subject. Although our modification does not allow to chose the type of weighting function, and does not have more fitting parameters, it fits the data at least as good as the original PT model, compare Table 3.

It is not surprising that the new model leads to different results on the parameters. – Even in the case of lotteries with only two outcomes where we can replace NPT with a special variant of PT, as shown above, the "effective weighting function" differs from the classical one, and so do the measured parameters. However, there is a clear correlation between the PT and the NPT parameters, both for the weighting function parameter and for the value function parameter, as can be seen from Fig. 3.

In the data analysis, we noticed as an interesting side observation that NPT shows the most advantage over PT for people who have a small weighting function parameter. In Fig. 4 we show the difference between the goodness of fit for both models in percentage. An reason for this effect could be the irregular shape of the classical weighting function by Tversky and Kahneman for small parameters. (The function becomes even non-monotone for very small values, compare [10].)
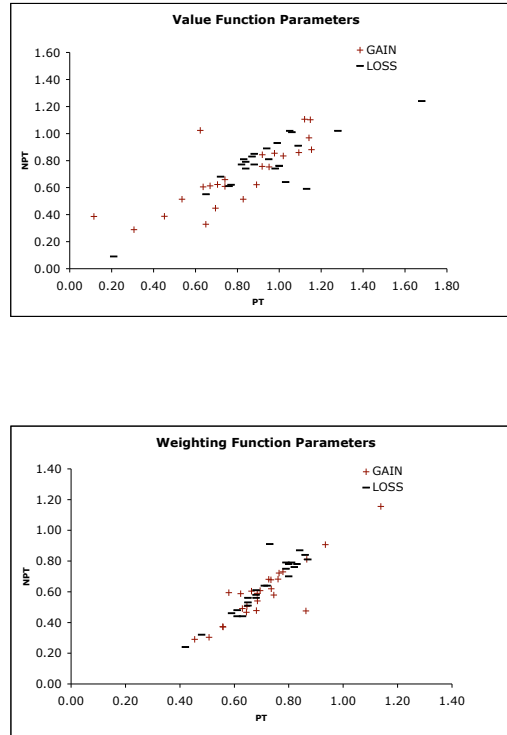
Figure 3: The correlation between the value (top) and weighting (bottom) parameters of PT and NPT for all test subjects. Data in gains is marked with a plus sign (+), data in losses with a minus sign (−).
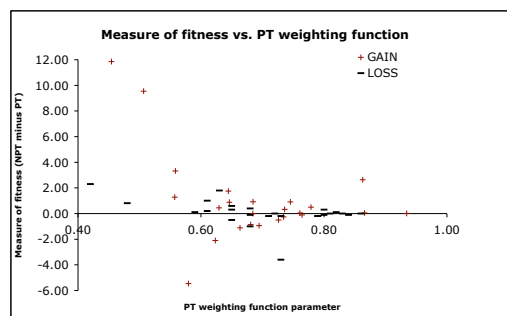

Figure 4: Difference of goodness of fit (R^2) between PT and NPT (a positive number indicates better fit of NPT) as a function of the weighting parameter in PT.

5 Conclusions

We have suggested two extensions of Prospect Theory: Normalized Prospect Theory (NPT) and Smooth Normalized Prospect Theory (SNPT). NPT allows to handle not only finite, but also infinite non-discrete lotteries. It shares the positive features with PT and fits experimental data even slightly better than classical PT, although it does not require more fitting parameters. One could argue that its weighting function is more naturally chosen than in PT which makes the goodness of fit even more remarkable. SNPT fixes the problem of continuity for PT in a natural way that resembles the "editing phase" by Kahneman and Tversky [7]. This model also eliminates many violations of stochastic dominance, although not all. Its limit for non-discrete lotteries corresponds to the limit of NPT.

With these improvements over the classical Prospect Theory, we allow for applications (e.g. in finance) of original PT that seemed up to now only possible with the conceptually different Cumulative Prospect Theory.

References

[1] Michael Birnbaum and T. Martin. Generalization across people, procedures, and predictions: Violations of stochastic dominance and coalescing. In *Emerging Perspectives on Decision Research*, pages 84–107. Cambridge University Press, New York, 2003.

[2] Michael Birnbaum and William. Ross. McIntosh. Violation of branch independence in choices between gambles. *Organizational Behavior and Human Decision Processes*, 67:91–110, 1996.

[3] Colin Camerer and Ho Teck-Hua. Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty*, 8:167–196, 1994.

[4] Hein Fennema and Peter P. Wakker. Original and new prospect theory: A discussion and empirical differences. *Journal of Behavioral Decision Making*, 10:53–64, 1997.

[5] Thorsten Hens, János Mayer, and Marc Oliver Rieger. From data to lotteries. in preparation.

[6] Steven. J. Humphrey. Event splitting effects or regret-aversion: More evidence under risk and uncertainty. *Journal of Risk and Uncertainty*, 11:263–274, 1995.

[7] Daniel Kahneman and Amos Tversky. Prospect Theory: An analysis of decision under risk. *Econometrica*, 47:263–291, 1979.

[8] R. Duncan Luce. Coalescing, event commutativity, and theories of utility. *Journal of Risk and Uncertainty*, 16:87–114, 1998.

[9] Robert J. McCann. Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J.*, 80(2):309–323, 1995.

[10] Marc Oliver Rieger and Mei Wang. Cumulative Prospect Theory and the St. Petersburg paradox. *Economic Theory*, 28:665–679, 2006.

[11] Chris Starmer. Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, 38:332–382, 2000.

[12] Chris Starmer and Robert Sugden. Testing for juxtaposition and event-splitting effects. *Journal of Risk and Uncertainty*, 6:235–254, 1993.

[13] Amos Tversky and Daniel Kahneman. Advances in Prospect Theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5:297–323, 1992.

[14] George Wu. An empirical test of ordinal independence. *Journal of Risk and Uncertainty*, 16:115–139, 1994.

[15] George Wu and Richard Gonzalez. Curvature of the probability weighting function. *Management Science*, 42:1676–1690, 1996.

[16] George Wu, Jiao Zhang, and Mohammed Abdellaoui. Testing prospect theories using probability tradeoff consistency. *Journal of Risk and Uncertainty*, 30(2):107–131, 2005.