European Journal of Operational Research 255 (2016) 103–109

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Stochastics and Statistics

Probability weighting and L-moments[☆]

Pavlo Blavatskyy[a,b,*]

[a] School of Business and Governance, Murdoch University, 90 South Street, Murdoch, WA 6150, Australia
[b] Montpellier Business School, Montpellier Research in Management, 2300 Avenue des Moulins, 34080 Montpellier, France

ARTICLE INFO

Article history:
Received 7 December 2015
Accepted 6 May 2016
Available online 31 May 2016

Keywords:
Decision under risk
Probability weighting function
Cumulative prospect theory
Mean-Gini approach
Theory of L-moments

ABSTRACT

Several popular generalizations of expected utility theory—cumulative prospect theory, rank-dependent utility and Yaari's dual model—allow for non-linear transformation of (de-)cumulative probabilities. This paper shows an unexpected connection between probability weighting and the statistical theory of L-moments. Specifically, cubic probability weighting results in a linear tradeoff between the expected value (the first L-moment), Gini (1912) mean difference statistic (the second L-moment, also known as L-scale) and the third L-moment (measuring skewness). Inverse S-shaped probability weighting function crossing the 45° line at a probability ≤0.5 reflects an aversion to the dispersion of outcomes and an attraction to positively skewed distributions.

1. Probability weighting and L-moments

The Allais (1953) paradox highlighted descriptive limitations of expected utility theory—people may reveal a different preference ordering over two pairs of probability distributions that must be ranked consistently by any expected utility maximizer. In response to the Allais (1953) paradox, expected utility theory was generalized to numerous non-expected utility theories (reviewed in Starmer, 2000). Popular generalizations of expected utility theory that can rationalize several behavioral regularities in choice under risk/uncertainty are Tversky and Kahneman (1992) cumulative prospect theory, [1] Quiggin (1981) rank-dependent utility and Yaari (1987) dual model. These theories introduce a non-linear probability weighting function over (de-)cumulative probabilities in choice under risk (or non-additive capacities over events in choice under uncertainty/ambiguity).

This paper shows an unexpected connection between probability weighting and the statistical theory of L-moments. Under Yaari (1987) dual model with a cubic probability weighting function preferences are represented by a weighted sum of three statistical measures: (1) the expected value of a lottery (which is also the first L-moment); (2) Gini (1912) mean difference statistic[2] (or the second L-moment, which is sometimes called L-scale); and (3) the third L-moment of a lottery (a measure of skewness). Thus, there is an unexpected connection to the financial literature.

Markowitz (1952) assumed that investor's preferences depend not only on the expected value (the mean) but also on the standard deviation (or the variance) of assets' returns. Unfortunately, any investor with such preferences inevitably violates the first-order stochastic dominance (*cf.* Borch 1969). Yitzhaki (1982) showed that violations of stochastic dominance can be avoided by using a different measure of statistical dispersion of assets' returns—Gini (1912) mean difference statistic.[3]

The mean-Gini approach of Shalit and Yitzhaki (1984) can be further extended by introducing a preference for gambling. Already Markowitz (1952, p. 90) considered the possibility that investors may care not only about the mean and the standard deviation (or the variance) but also—about the skewness of assets' returns. Yet, Markowitz (1952, p. 90) proposed to measure skewness with the third central moment, which may lead to the violations of the first-order stochastic dominance. Such violations may be avoided by

☆ I am grateful to the participants of Risk, Uncertainty and Decision (RUD) 2014 conference at the University of Warwick for helpful comments.

* Correspondence address: School of Business and Governance, Murdoch University , Australia. Tel.: +61 473646841, +33 6 03 15 83 42; fax: +61 89360 6966
E-mail address: p.blavatskyy@montpellier-bs.com, pavlo.blavatskyy@gmail.com

[1] For example, cumulative prospect theory can account for Allais (1953) common consequence effect, the common ratio effect (*e.g.* Bernasconi, 1994), systematic violations of the betweenness axiom (*e.g.* Camerer and Ho, 1994) and the four-fold pattern of risk attitudes. Schmidt, Starmer, and Sugden (2008) present an extension of the theory that can account for the preference reversal phenomenon as well as the discrepancy between willingness-to-accept and willingness-to-pay. Yet, there are also several behavioral regularities that cumulative prospect theory fails to rationalize such as Blavatskyy (2012b) troika paradox and Machina (2009) reflexion example (see also Blavatskyy, 2013a). Curiously, typical parameterizations of the theory cannot resolve the classical St. Petersburg paradox (Blavatskyy, 2005).

[2] Mathematical expectation of the absolute value of the difference between two realizations of a lottery.

[3] Blavatskyy (2010b) showed that violations of stochastic dominance can be also avoided by measuring financial risks with the mean absolute semideviation (*i.e.* by aggregating only those deviations that are below the expected value).

http://dx.doi.org/10.1016/j.ejor.2016.05.007
0377-2217/

using the third L-moment (Hosking, 1990) instead of the third central moment. L-moments are more robust than conventional moments to outliers, which turns out to be sufficient for ruling out violations of stochastic dominance. Note that the first L-moment is the expected value and the second L-moment is simply one half of Gini (1912) mean difference statistic.

The literature on financial decision making uses the model of multi-attribute choice (with attributes being different moments of the distribution of assets' returns). Arguably the simplest decision criterion in multi-attribute choice is to aggregate different attributes into one real-valued index. Such a linear trade-off between the first three L-moments of a lottery represents preferences under Yaari (1987) dual model with a cubic probability weighting function.

A decision maker who prefers positively skewed distributions (*e.g.*, a small chance to win a highly desirable outcome) and dislikes negatively skewed distributions (*e.g.*, a small chance to end up with a highly undesirable outcome) generally has an inverse S-shaped probability weighting function. Moreover, this function crosses the 45° line at a probability smaller (greater) than 0.5 if a decision maker is also averse (attracted) to the dispersion of outcomes. Thus, a well-known probability weighting function empirically discovered by Tversky and Kahneman (1992, p. 309) can be intuitively rationalized as a combination of two factors: an aversion to the second L-moment (dispersion) and an attraction to the third L-moment (skewness). A decision maker not caring about skewness has a simpler probability weighting function—a quadratic polynomial of (de-)cumulative probabilities—that can only be either concave or convex. This probability weighting function is discussed in Delquié and Cillo (2006, pp. 204–205). Yaari (1987) dual model with such a probability weighting function is a special case of the mean-Gini approach of Shalit and Yitzhaki (1984).

Economic decision theory deviated from the idea of risk neutrality by introducing a non-linear (Bernoulli) utility function over money as well as a non-linear probability weighting function over (de-) cumulative probabilities. Financial decision theory deviated from the same idea by introducing a preference for the higher moments of a probability distribution. This paper shows that representing preferences with the first three L-moments is *de facto* equivalent to introducing a cubic probability weighting function. Thus, the two complimentary approaches to modeling decision making under risk from economics and finance can be unified into one general theory.

The remainder of the paper is structured as follows. Cumulative prospect theory is briefly sum-marized in Section 2. Readers familiar with the topic may skip Section 2 without the loss of continuity. A cubic probability weighting function is presented in Section 3. Its relation to statistical L-moments (Hosking, 1990) and mean-Gini approach (Shalit & Yitzhaki, 1984) is discussed in Section 4. Section 5 concludes with a general discussion.

2. Cumulative prospect theory for choice under risk

Let $X \subseteq \mathbb{R}$ denote a nonempty set of possible outcomes (*e.g.* financial returns). A lottery $L{:}X{\to}[0,1]$ is a discrete probability distribution on set X, *i.e.*, $L(x)\epsilon[0,1]$ for all $x \in X$ and $\Sigma_{x\epsilon X}L(x) = 1$. Any lottery can be alternatively characterized by its cumulative distribution function $F_L{:}X{\to}[0,1]$. This function gives the probability that lottery L yields an outcome at most as good as outcome $x \in X$:

$$F_L(x) = \sum_{y \in X, x \geq y} L(y) \tag{1}$$

A lottery can be also characterized by its decumulative distribution function $G_L{:}X{\to}[0,1]$. This function gives the probability that lottery L yields an outcome at least as good as outcome $x \in X$:

$$G_L(x) = 1 - F_L(x) + L(x) \tag{2}$$

In cumulative prospect theory one outcome $r \in X$ is the reference point of a decision maker. Outcomes greater than the reference point are called gains. The set of all gains is denoted by $X_+ \subseteq X$. Outcomes smaller than the reference point are called losses. The set of all losses is denoted by $X_- \subseteq X$.

Preferences of a decision maker are represented by the following utility function:

$$U(L) = \sum_{x \in X_-} [w_-(F_L(x)) - w_-(1 - G_L(x))]u(x) + \sum_{x \in X_+} [w_+(G_L(x)) - w_+(1 - F_L(x))]u(x) \tag{3}$$

where $w_-[0,1]{:}{\to}[0,1]$ and $w_+[0,1]{:}{\to}[0,1]$ are two strictly increasing probability weighting functions such that $w_-(0) = w_+(0) = 0$ and $w_-(1) = w_+(1) = 1$; and $u{:}X{\to}\mathbb{R}$ is an increasing utility function that is unique up to a multiplication by a positive constant and satisfying $u(r) = 0$.[4]

Quiggin (1981) rank-dependent utility is a special case of cumulative prospect theory when either $w_-(p) = 1 - w_+(1-p)$ for all $p \in [0,1]$ or all outcomes in X are greater then the reference point r (so that the set of losses X_- is empty). Yaari (1987) dual model is a special case of rank-dependent utility when utility function is linear: $u(x) = x$ for all $x \in X$. Expected utility theory is a special case of rank-dependent utility when a probability weighting function is linear: $w_+(p) = p$ for all $p \in [0,1]$.

3. A cubic probability weighting function

In the following, a probability weighting function is written without subscripts "+" and "−" whenever it is inconsequential whether we deal with gains or losses. We consider a probability weighting function that is a cubic polynomial of probability:

$$w(q) = q - \rho \cdot q(1-q) + \tau \cdot q(1-q)(1-2q) \tag{4}$$

for all $q \in [0,1]$ and two subjective parameters $\rho, \tau \in \mathbb{R}$. Note that function (4) always satisfies $w(0) = 0$ and $w(1) = 1$. Table 1 summarizes the properties of function (4) for various values of parameters ρ and τ.

Fig. 1 plots function (4) for several positive values of parameter τ. Note that the probability weighting function is inverse S-shaped crossing the 45° line at a probability q less than one half when ρ is positive but less then τ (*cf.* dashed curves in Fig. 1). Yet, if ρ is greater than or equal to τ, the probability weighting function does not cross the 45° line at all (*cf.* a solid curve in Fig. 1). When ρ is negative but greater than $-\tau$, the probability weighting function is inverse S-shaped crossing the 45° line at a probability q greater than one half (*cf.* dotted and dashed-dotted curves in Fig. 1). Yet, if ρ is less than or equal to $-\tau$, the probability weighting function does not cross the 45° line at all (*cf.* a dashed-double-dotted curve in Fig. 1). Thus, probability weighting function (4) with a positive value of τ is quite flexible. It can take a variety of shapes including a convex function, a concave function and an inverse S-shaped function crossing the 45° line at various probabilities q.

We can use an existing axiomatization of cumulative prospect theory (with a generic probability weighting function) and impose

[4] There are also additional convexity assumptions. Tversky and Kahneman (1992, p. 305) assumed that both probability weighting functions are inverse S-shaped (concave near probability zero and convex near probability one). This paper relaxes this assumption. Additionally, prospect theory assumes that utility function is convex on X_- and concave on X_+. Finally, the assumption of loss aversion restricts utility function as well (see Köbberling and Wakker, 2005; Blavatskyy, 2011b).

Table 1
Properties of function (4) for various values of parameters ρ and τ.

Values of parameters ρ and/or τ	Function (4) is monotone when	Function (4) crosses the 45° line at	Shape of function (4)
$\rho = \tau = 0$	Always	Coincides with 45° line	Linear
$\tau = 0$[a], $\rho<0$	$-1\leq\rho<0$	Never	Globally concave
$\tau = 0$[a], $\rho>0$	$0<\rho\leq 1$	Never	Globally convex
$\rho = 0$, $\tau<0$	$-1\leq\tau<0$	$q = ½$	S-shaped
$\rho = 0$, $\tau>0$	$0<\tau\leq 2$	$q = ½$	Inverse S-shaped
$0<\tau\leq 2$, $-\tau<\rho<\tau$	$\lvert\rho\rvert\leq 1+\tau$ when $\tau\leq ½$; $\rho^2\leq 3\tau(2-\tau)$ when $\tau\geq ½$	$q = ½(1-\rho/\tau)$	Inverse S-shaped
$0<\tau\leq 2$, $\rho\geq\tau$	$\lvert\rho\rvert\leq 1+\tau$ when $\tau\leq ½$; $\rho^2\leq 3\tau(2-\tau)$ when $\tau\geq ½$	Never	Inverse S-shaped/convex
$0<\tau\leq 2$, $\rho\leq-\tau$	$\lvert\rho\rvert\leq 1+\tau$ when $\tau\leq ½$; $\rho^2\leq 3\tau(2-\tau)$ when $\tau\geq ½$	Never	Inverse S-shaped/concave

[a] This quadratic probability weighting function is discussed in Delquié and Cillo (2006, pp. 204–205).

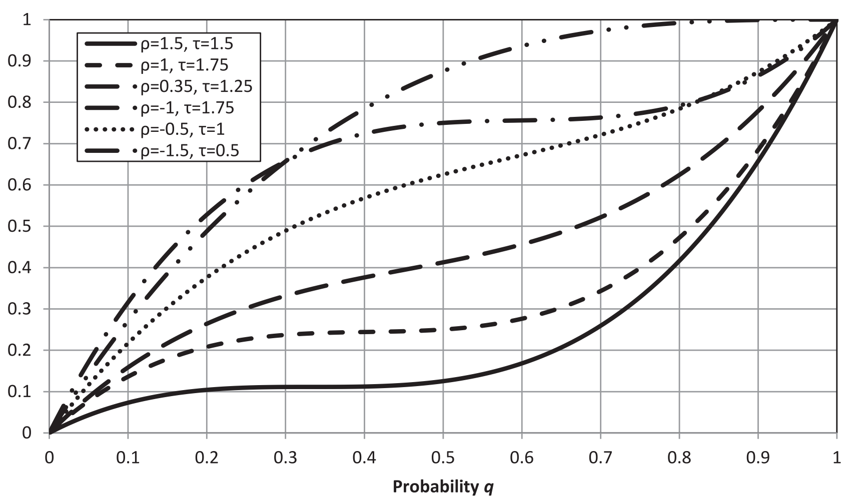

Fig. 1. Probability weighting function (4) when $\tau>0$.

an additional restriction on preferences to obtain the utility function of cumulative prospect theory with a cubic probability weighting function (4). This approach is illustrated in the appendix by imposing a restriction on the standard sequence of probability equivalents (Abdellaoui, 2000, p. 1501).

4. Statistical theory of L-moments

The first L-moment $\lambda_1(L)$ of lottery L is the same as the first central moment (the expected value):

$$\lambda_1(L) = \sum_{x\in X} L(x)\cdot x \tag{5}$$

Note that we can alternatively rewrite the definition of the first L-moment (5) as follows:

$$\lambda_1(L) = \sum_{x\in X} [w_1(G_L(x)) - w_1(1 - F_L(x))]\cdot x \tag{6}$$

where the probability weighting function w_1:[0,1]→[0,1] is given by

$$w_1(q) = q \tag{7}$$

Formula (6) simply states that the first L-moment of distribution L is nothing but Quiggin (1981) rank-dependent utility of L with a linear probability weighting function (7) and a linear utility function (*i.e.*, utility of L in Yaari (1987) dual model with a linear probability weighting function (7)).

The second L-moment, also known as L-scale, measures the dispersion of lottery's outcomes. Intuitively, the second L-moment can be explained as follows. Consider an investor who invested in asset L in two periods. The expected difference between the higher and the lower return of the two realized outcomes characterizes the dispersion of L's outcomes. If outcomes are clustered close to each other then this expected difference is small (and it is zero if and only if L yields one outcome for certain). If outcomes are scattered far apart then this expected difference is large. Formally, the second L-moment $\lambda_2(L)$ of a discrete probability distribution L is defined as (8).

$$\lambda_2(L) = \sum_{x\in X} L(x)\cdot [F_L(x) - G_L(x)]\cdot x \tag{8}$$

An alternative definition of the second L-moment that will be useful in the context of this paper is

$$\lambda_2(L) = \sum_{x\in X} [w_2(G_L(x)) - w_2(1 - F_L(x))]\cdot x \tag{9}$$

where function w_2:[0,1]→[0,¼] is defined as

$$w_2(q) = q\cdot(1-q) \tag{10}$$

According to definition (9), we can think of the second L-moment of distribution L as Quiggin (1981) rank-dependent utility of L with a quadratic "probability weighting" function (10) and a linear utility function (or as utility of L in Yaari's dual model with a quadratic "probability weighting" function (10)).

The second L-moment is one half of Gini (1912) mean difference statistic. If a discrete probability distribution L yields only positive returns ($X \subseteq \mathbb{R}_+$) then the ratio $\lambda_2(L)/\lambda_1(L)$ is equal to Gini's coefficient.

The third L-moment measures the skewness of outcomes and it can be explained as follows. Consider an investor who invested in an asset in three periods. The expected difference between the mean and the median return of the three realized outcomes characterizes the skewness of the distribution. For a symmetric distribution this expected difference is zero. For a positively (negatively) skewed distribution this expected difference is positive (negative).

Formally, the third L-moment $\lambda_3(L)$ is

$$\sum_{x \in X} L(x) \cdot [(F_L(x) - G_L(x))^2 - F_L(x)(1 - F_L(x)) - G_L(x)(1 - G_L(x))] \cdot x \quad (11)$$

An alternative definition of the third L-moment that will be useful in the context of this paper is

$$\lambda_3(L) = \sum_{x \in X} [w_3(G_L(x)) - w_3(1 - F_L(x))] \cdot x \quad (12)$$

where function w_3:[0,1]→[$-\sqrt{3}/18$,$\sqrt{3}/18$] is defined as

$$w_3(q) = q \cdot (1 - q) \cdot (1 - 2q) \quad (13)$$

According to definition (12), the third L-moment of distribution L is nothing but Quiggin (1981) rank-dependent utility of L with a cubic "probability weighting" function (13) and a linear utility function (or utility of L in Yaari (1987) dual model with a cubic "probability weighting" function (13)).

In general, the r-th L-moment of lottery L, $r \in \mathbb{N}$, is given by

$$\lambda_r(L) = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E\left(L_{r-k,r}\right) \quad (14)$$

where $E(L_{r-k,r})$ is the expected value of the (r–k)th ranked realized outcome of lottery L (also known as (r–k)th order statistic) when L is played out r times (*cf.* Hosking, 1990, p. 106, Eq. (2.1)). Since moments (14) are expectations of **L**inear combinations of order statistics, they are called L-moments.

Using definitions (6), (9) and (12) we immediately obtain the following result. Utility function of Yaari (1987) dual model[5] with our proposed probability weighting function (4) is a weighted sum of the first three L-moments:

$$\lambda_1(L) - \rho \cdot \lambda_2(L) + \tau \cdot \lambda_3(L) = \sum_{x \in X} [w(G_L(x)) - w(1 - F_L(x))] \cdot x \quad (15)$$

Thus, two subjective parameters ρ and τ of our proposed probability weighting function can be interpreted as follows. Parameter ρ captures a decision maker's attitude to the dispersion of outcomes and parameter τ—to the skewness of outcomes. We expect that a typical decision maker dislikes choice alternatives yielding widely dispersed or negatively skewed outcomes. Therefore, the second L-moment enters with a negative sign and the third L-moment—with a positive sign[6] into utility (15).

Eq. (15) demonstrates that our proposed probability weighting function is related to Shalit and Yitzhaki (1984) mean-Gini approach to optimal portfolio investment. Shalit and Yitzhaki (1984) considered investors' preferences that depend only on the expected return and Gini (1912) mean difference statistic. Since the first L-moment is the expected return and the second L-moment is one half of Gini (1912) mean difference statistic, the preferences considered in Shalit and Yitzhaki (1984) are represented by utility function $U(\lambda_1(L), \lambda_2(L))$ that is increasing in the first argument and decreasing in the second argument. A special case of such utility function is utility function (15) that is independent of the third argument (L-skewness), *i.e.* when $\tau = 0$.

Sometimes it is convenient to consider a continuous probability distribution. For a continuous random variable M on $\mathbb{R}$ with a cumulative distribution function F_M:X→[0,1] the first L-moment is

$$\lambda_1(M) = \int_{-\infty}^{+\infty} x dF_M(x) \quad (16)$$

the second L-moment is

$$\lambda_2(M) = \int_{-\infty}^{+\infty} x[2F_M(x) - 1]dF_M(x) = \int_{-\infty}^{+\infty} x dF_M^2(x) - \lambda_1(M) \quad (17)$$

and the third L-moment is

$$\lambda_3(M) = \int_{-\infty}^{+\infty} x\left[6F_M^2(x) - 6F_M(x) + 1\right]dF_M(x) = 2\int_{-\infty}^{+\infty} x dF_M^3(x) - 3\lambda_2(M) - 2\lambda_1(M) \quad (18)$$

5. Conclusion

This paper considers a cubic probability weighting function for cumulative prospect theory (and its special cases—rank dependent utility and Yaari's dual model). With only two subjective parameters this function can nonetheless take a variety of forms including a concave/convex shape, an inverse S-shape and an S-shape. Such a function is useful in estimating risk preferences from a heterogeneous population. Our reexamination of two empirical data sets shows that cubic function compares favorably to the existing probability weighting functions. A comparison with a nonparametric probability weighting function shows that we sacrifice relatively little in terms of goodness of fit to the data for a parsimonious two-parameter functional form.

Beyond its practical utility in applied microeconomics the proposed function contributes to the intuitive understanding of cumulative prospect theory. The original prospect theory (Kahneman & Tversky, 1979) transformed probabilities of lottery's outcomes. In contrast, the probability weighting function in cumulative prospect theory transforms (de-)cumulative probabilities of outcomes. The transformation of (de-)cumulative probabilities arguably does not have an immediate intuitive appeal (*e.g.*, Luce, 1996, p. 85; see, however, Diecidue & Wakker, 2001). A characteristic behavioral property of cumulative prospect theory—a sign-comonotonic tradeoff consistency[7]—is a complex axiom that might seem not immediately convincing (*e.g.*, Safra & Segal, 1998, p. 29).

Yet, a special case of cumulative prospect theory—Yaari's dual model with a cubic probability weighting function—is equivalent to a linear tradeoff between the expected value, a measure of statistical dispersion (Gini's statistic or L-scale) and a measure of skewness (the third L-moment). Thus, a non-linear transformation of (de-)cumulative probabilities with an inverse S-shaped probability weighting function has a rather intuitive explanation. It results from a combination of two behavioral forces: an aversion to the dispersion of lottery's outcomes and a preference for gambling (an attraction to positively skewed distributions and an aversion to negatively skewed distributions).

A tradeoff between the expected value and dispersion (a risk measure) is one of the cornerstones in the theory of optimal portfolio investment. Thus, a cubic probability weighting function builds an unexpected bridge between a microeconomic and a financial decision theory. Both approaches to decision making under risk can be unified under one general theory—a cumulative prospect theory with a cubic probability weighting function. We can think of an economic and a financial decision theory as the two sides of the same coin. Positive synergy effects are apparent for both disciplines. Economic decision theory receives a new intuition for non-linear probability weighting whereas financial decision theory benefits from a solid behavioral characterization of cumulative prospect theory.

[5] That is, Quiggin (1981) rank-dependent utility with a linear utility function.

[6] Ebert (2015, Section 5.5.1, p. 94) and Astebro et al. (2015, Section 3.1, p. 198) report experimental evidence of skewness-seeking choices.

[7] Blavatskyy (2013b) recently showed that tradeoff consistency can be weakened to cardinal independence, which is also known as standard sequence invariance.

A flexible weighting function helps in understanding the properties of prospect theory. For example, a decision maker with an inverse S-shaped function (4) that crosses the 45° line at a probability greater than one half can simultaneously exhibit the common ratio effect in classical common ratio problems[8] and the reverse common ratio effect—in another type of problems[9] (*cf.* Blavatskyy, 2010a, p. 222, pair #6 in Table 1). This property of cumulative prospect theory remained previously unnoticed in the literature.

Function (4) fits well to the experimental data collected by Hey and Orme (1994, p. 1293).[10] Median maximum likelihood estimates of its parameters are $\rho = 0.02$, $\tau = 0.24$ ($\rho = 0.02$, $\tau = 0.29$) using Blavatskyy (2014) model of random errors drawn from the normal (beta) distribution. Most subjects reveal positive τ, *i.e.* they are prone to gambling. Half of the subjects reveal positive ρ, *i.e.* there is no strong aversion to outcome dispersion. A typical subject has an inverse S-shaped weighting function but the point of crossing the 45° line is about equally likely to be above or below 0.5. Function (4) apparently outperforms the power and Prelec (1998) weighting functions but it is comparable to those proposed by Tversky and Kahneman (1992, p. 309) and Goldstein and Einhorn (1987).

Function (4) also fits well to data collected by Blavatskyy (2013c). Replacing a nonparametric probability weighting function with a two-parameter function (4) significantly reduces the goodness of fit for about 5–10% of all subjects. In contrast, by using even more parsimonious one-parameter functions proposed by Tversky and Kahneman (1992) and Prelec (1998) we significantly reduce the goodness of fit for about one third and one quarter of all subjects respectively.[11]

Appendix

Behavioral characterization of cubic probability weighting function (4)

We investigate how the behavior of a decision maker is restricted when his or her probability weighting function takes on our proposed functional form (4). Let S denote a non-empty set of the states of the world. Only one state of the world is true but a decision maker does not know which one. There is a sigma-algebra Σ of the subsets of S that are called events. Let X be a connected set of outcomes. An act $f{:}S\to X$ is a Σ-measurable function from S to X. The set of all acts is denoted by $\mathcal{F}$. A constant act that yields the same outcome $x \in X$ in all states of the world is denoted by $\boldsymbol{x}\in\mathcal{F}$. For any partition $\{E_1,\ldots,E_n\}$ of the state space S into n events[12] let $\{x_1,E_1;\ \ldots\ ;x_n,E_n\}$ denote a step act that yields outcome $x_i\in X$ in any state $s\in E_i$, for $i\in\{1,\ \ldots\ ,n\}$. Let $\mathbb{F}\subset\mathcal{F}$ denote the set of all step acts. For compact notation, let $xEf\in\mathbb{F}$ denote a step act that yields an outcome $x \in X$ in all states $s\in E$, $E\in\Sigma$, and an outcome $f(s)$ in the remaining states $s\in S\backslash E$. An event $E \in \Sigma$ is null (or inessential) if $xEf\succeq yEf$ for all $x, y\in X$ and $f{:}S\backslash E\to X$. Otherwise, the event is nonnull (or essential). If there is only one nonnull event, we additionally assume that X is a separable set.

A decision maker has a *preference relation* $\succeq$ on $\mathcal{F}$. As usual, the symmetric part of $\succeq$ is denoted by $\sim$ and the asymmetric part of $\succeq$ is denoted by $\succ$. To begin with, we assume that a decision maker behaves as if he or she maximizes the utility function of cumulative prospect theory. The necessary and sufficient axioms for such a preference representation are listed below.

Axiom 1 (*Completeness*) For all $f, g\in\mathcal{F}$ either $f\succeq g$ or $g\succeq f$ (or both).

Axiom 2 (*Transitivity*) For all $f, g, h\in\mathcal{F}$ if $f\succeq g$ and $g\succeq h$ then $f\succeq h$.

Next, we assume either continuity of preferences (Axiom 3 below), which is known as the connected topology approach, or solvability and Archemidean property (Axioms 4 and 5 below), which is known as the algebraic approach.[13]

Axiom 3 (*Step-continuity*) For any partition $\{E_1,\ \ldots\ ,E_n\}$ of set S into n events and any step act $\{x_1,E_1;\ \ldots\ ;x_n,E_n\}$ the sets $\{(y_1,\ \ldots\ ,y_n)\in X^n{:}\{y_1,E_1;\ \ldots\ ;y_n,E_n\}\succeq\{x_1,E_1;\ \ldots\ ;x_n,E_n\}\}$ and $\{(y_1,\ \ldots\ ,y_n)\in X : \{x_1,E_1;\ \ldots\ ;x_n,E_n\}\succeq\{y_1,E_1;\ \ldots\ ;y_n,E_n\}\}$ are closed with respect to the product topology on X^n.

Axiom 4 (*Solvability*) For all $x, y\in X$, $E\in\Sigma$, $f{:}S\backslash E\to X$ and $g\in\mathbb{F}$ such that $xEf\succeq g\succeq yEf$ there exists an outcome $z\in X$ such that $g\sim zEf$.

Axiom 5 (*Archimedean axiom*) A sequence of outcomes $\{x_i\}_{i\in\mathbb{N}}$ such that $x_iEg\sim x_{i-1}Ef$ and $yEf\succeq x_iEf\succeq xEf$ for some $x, y\in X$ is finite for all $x_0\in X$, non-null $E\in\Sigma$, and $f, g{:}S\backslash E\to X$ such that either $x_0Ef\succ x_0Eg$ or $x_0Eg\succ x_0Ef$.

Cumulative prospect theory is traditionally characterized by sign-comonotonic tradeoff consistency (Tversky & Kahneman, 1992, p. 319). Blavatskyy (2013b) recently showed that tradeoff consistency can be further weakened to an axiom known as cardinal independence or standard sequence invariance (Krantz, Luce, Suppes, & Tversky, 1971, Section 6.11.2). We use this weaker Axiom 6.

Axiom 6 (*Sign-comonotonic Cardinal Independence*) If $xEf\succeq yEg$, $xEg\succeq zEf$, and $yAh\succeq xAk$ then $xAh\succeq zAk$ for all $x, y, z \in X$; $f, g{:}S\backslash E\to X$; $h, k{:}S\backslash A\to X$, any nonnull event $E\in\Sigma$ and any event $A\in\Sigma$, provided that all acts are pairwise comonotonic[14] and outcomes are either all gains or all losses.

We shall now characterize a probability weighting function for gains (the argument is analoguous for losses). Consider any standard sequence $\{x_i\}_{i\in\{1,\ \ldots\ ,\ m\}}$ of $m\geq 3$ gain outcomes such that $x_iEy\sim x_{i-1}Ez$ for some non-null event $E\in\Sigma$ and outcomes $y,z,x_0\in X_+$ such that $\boldsymbol{x_0}\succ\boldsymbol{y}$, $\boldsymbol{x_0}\succ\boldsymbol{z}$ and $x_0Ez\succ x_0Ey$. If Axioms 1, 2, 6 and either 3 or 4, 5 hold then any such sequence contains outcomes that are equally spaced on the utility scale (*cf.* Wakker and Deneffe, 1996, p. 1144). For any standard sequence $\{x_i\}_{i\in\{1,\ \ldots\ ,\ m\}}$ of outcomes we can construct a standard sequence of probability equivalents $\{p_i\}_{i\in\{1,\ \ldots\ ,\ m\}}$ such that $\boldsymbol{x_i}\sim x_mE_ix_0$ where the objective probability of event E_i is $p_i \in [0,1]$. If Axioms 1, 2, 6 and either 3 or 4, 5 hold then $w_+(p_i) = i/m$ (*e.g.* Abdellaoui, 2000, p. 1501). In other words, a standard sequence of m probability equivalents is the inverse of the probability weighting function over values i/m, $i \in \{1,\ \ldots\ ,m\}$.

In cumulative prospect theory the standard sequence of probability equivalents can be any monotonically increasing sequence of probabilities. Yet, if a probability weighting function of cumulative prospect theory takes on a specific form (4), the standard sequence of probability equivalents is restricted by the following Axiom 7.

---

[8] In the terminology of Loomes and Sugden (1998), these are the common ratio problems with a gradient greater than one. Examples of such problems can be found in Ballinger and Wilcox (1997) and Blavatskyy (2010a, p. 234, pair #7 in Table 3).

[9] In the terminology of Loomes and Sugden (1998), these are the common ratio problems with a gradient less than one.

[10] This well-known data set is often reexamined in the literature, *cf.* Hey (1995), Hey and Carbone (1995), Carbone and Hey (1995), Buschena and Zilberman (2000), Blavatskyy (2007), Wilcox (2008, 2010).

[11] These results appear to be robust to different models of probabilistic choice (*cf.* Blavatskyy and Pogrebna, 2010). We considered Fechner (1860) model with normally distributed random errors (cf. Blavatskyy, 2008), Fishburn (1978, p. 635) model with a power function $\rho(v) = v^{\mu}$ for all $v\geq 0$ (μ is a subjective "noise" parameter) and Blavatskyy (2009, 2011a, 2012a) model with function $\varphi(v) = e^{\lambda v}-1$ for all $v\geq 0$ (λ is a subjective "noise" parameter).

[12] That is, $E_i\in\Sigma$ for all $i \in\{1,\ldots,n\}$, $E_1\cup\ldots\cup E_n=S$ and $E_i\cap E_j = \emptyset$ for all $i,j\in\{1,\ldots,n\}$, $i\neq j$.

[13] See Section 6.11 in Kranzt et al. (1971, p. 301), Wakker (1988), Köbberling and Wakker (2003, p. 398).

[14] Acts $f,g \in \mathcal{F}$ are comonotonic if there are no states of the world $s,t \in S$ such that $\boldsymbol{f(s)}\succ\boldsymbol{f(t)}$ but $\boldsymbol{g(t)}\succ\boldsymbol{g(s)}$.

Axiom 7 For any standard sequence of $m \geq 3$ probability equivalents $\{p_i\}_{i \in \{1, \ldots, m\}}$ the ratio

$$\frac{\frac{p_j(p_i - i/m)}{p_i(1-p_i)} - \frac{p_i(p_j - j/m)}{p_j(1-p_j)}}{p_i - p_j}$$

and the ratio

$$\frac{\frac{p_i - i/m}{p_i(1-p_i)} - \frac{p_j - j/m}{p_j(1-p_j)}}{p_i - p_j}$$

are constant for all $i, j \in \{1, \ldots, m-1\}$, $i \neq j$.

For an expected utility maximizer the probability equivalent p_i is simply equal to i/m for all $i \in \{1, \ldots, m\}$. In this case, both ratios in Axiom 7 are equal to zero and Axiom 7 holds trivially. More generally, for a decision maker, who behaves as if maximizing utility function (3) with a probability weighting function (4) for gains, the first ratio in Axiom 7 is equal to $\tau - \rho$ and the second ratio in Axiom 7 is equal to 2τ.[15] In other words, both ratios are constant for a cubic probability weighting function.

Axiom 7 is not only necessary for a probability weighting function (4) but it is also sufficient. Note that if Axiom 7 holds then any probability equivalent p_i from a standard sequence of probability equivalents $\{p_i\}_{i \in \{1, \ldots, m\}}$ must satisfy the equation $p_i - \rho p_i(1-p_i) + \tau p_i(1-p_i)(1-2p_i) = i/m$ for some constant $\rho, \tau \in \mathbb{R}$. By construction, any probability equivalent p_i from a standard sequence is such that $w(p_i) = i/m$. Thus, if Axiom 7 holds then the probability weighting function takes on functional form (4) at any probability value that belongs to a standard sequence of probability equivalents.

Let us now consider an event E occurring with an objective probability q that does not belong to a standard sequence of probability equivalents $\{p_i\}_{i \in \{1, \ldots, m\}}$ for any $m \geq 3$. Let $\underline{E}$ denote an event occurring with the highest objective probability $\underline{p}$ that belongs to the standard sequence of probability equivalents $\{p_i\}_{i \in \{1, \ldots, m\}}$ such that $x_m E x_0 \succ x_m \underline{E} x_0$. If no such event $\underline{E}$ exists we denote by $\underline{E}$ the null event $\emptyset$ occurring with objective probability zero. Similarly, let $\hat{E}$ denote an event occurring with the lowest objective probability $\overline{p}$ that belongs to the standard sequence of probability equivalents $\{p_i\}_{i \in \{1, \ldots, m\}}$ such that $x_m \hat{E} x_0 \succ x_m E x_0$. If no such event $\hat{E}$ exists we denote by $\hat{E}$ the universal event S occurring with objective probability one. Then the value of the probability weighting function at probability q is bounded by $w(\underline{p}) < w(q) < w(\overline{p})$. By construction, the difference between objective probabilities of events $\hat{E}$ and $\underline{E}$ is $\overline{p} - \underline{p} = 1/m$. Thus, if the number m of elements in the standard sequence of probability equivalents increases, the value $w(\underline{p})$ converges to the value $w(\overline{p})$. Hence, the value $w(q)$, which is bounded by $w(\underline{p})$ and $w(\overline{p})$ must also take on the functional form (4).

References

Abdellaoui, M. (2000). Parameter-free elicitation of utility and probability weighting functions. *Management Science, 46*, 1497–1512.
Allais, M. (1953). Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulates et Axiomes de l'Ecole Américaine. *Econometrica, 21*, 503–546.
Astebro, T., Mata, J., & Santos-Pinto, L. (2015). Skewness seeking: Risk loving, optimism or overweighting of small probabilities? *Theory and Decision, 78*, 189–208.
Ballinger, P., & Wilcox, N. (1997). Decisions, error and heterogeneity. *Economic Journal, 107*, 1090–1105.

[15] In this case the probability equivalent p_i is equal to the inverse of the weighting function at i/m: $p_i = w^{-1}(\frac{i}{m}) = \frac{1}{9\tau}\left(3\tau - \rho - \sqrt[3]{\frac{\Delta + \sqrt{\Delta^2 - 4(\rho^2 + 3\rho\tau - 9\tau)^3}}{2}} + \frac{9\tau - 3\rho\tau - \rho^2}{\sqrt[3]{\frac{\Delta + \sqrt{\Delta^2 - 4(\rho^2 + 3\rho\tau - 9\tau)^3}}{2}}}\right)$ where $\Delta = 2(\rho - 3\tau)^3 - 27\tau(\rho - 3\tau)(1 - \rho + \tau) + 243\tau^2(i/m)$.

Bernasconi, M. (1994). Nonlineal preference and two-stage lotteries: Theories and evidence. *Economic Journal, 104*, 54–70.
Blavatskyy, P., & Pogrebna, G. (2010). Models of stochastic choice and decision theories: Why both are important for analyzing decisions. *Journal of Applied Econometrics, 25*(6), 963–986.
Blavatskyy, P. (2005). Back to the St. Petersburg paradox? *Management Science, 51*(4), 677–678.
Blavatskyy, P. (2007). Stochastic expected utility theory. *Journal of Risk and Uncertainty, 34*(3), 259–286.
Blavatskyy, P. (2008). Stochastic utility theorem. *Journal of Mathematical Economics, 44*, 1049–1056.
Blavatskyy, P. (2009). Preference reversals and probabilistic choice. *Journal of Risk and Uncertainty, 39*(3), 237–250.
Blavatskyy, P. (2010a). Reverse common ratio effect. *Journal of Risk and Uncertainty, 40*, 219–241.
Blavatskyy, P. (2010b). Modifying the mean-variance approach to avoid violations of stochastic dominance. *Management Science, 56*(11), 2050–2057.
Blavatskyy, P. (2011a). A model of probabilistic choice satisfying first-order stochastic dominance. *Management Science, 57*(3), 542–548.
Blavatskyy, P. (2011b). Loss aversion. *Economic Theory, 46*(1), 127–148.
Blavatskyy, P. (2012a). Probabilistic choice and stochastic dominance. *Economic Theory, 50*(1), 59–83.
Blavatskyy, P. (2012b). Troika paradox. *Economics Letters, 115*(2), 236–239.
Blavatskyy, P. (2013a). Two examples of ambiguity aversion. *Economics Letters, 118*(1), 206–208.
Blavatskyy, P. (2013b). A simple behavioral characterization of subjective expected utility. *Operations Research, 61*(4), 932–940.
Blavatskyy, P. (2013c). Which decision theory? *Economics Letters, 120*(1), 40–44.
Blavatskyy, P. (2014). Stronger utility. *Theory and Decision, 76*(2), 265–286.
Borch, K. (1969). A note on uncertainty and indifference curves. *The Review of Economic Studies, 36*(1), 1–4.
Buschena, D., & Zilberman, D. (2000). Generalized expected utility, heteroscedastic error, and path dependence in risky choice. *Journal of Risk and Uncertainty, 20*, 67–88.
Camerer, F. C., & Ho, T.-H. (1994). Violations of the betweenness axiom and nonlinearity in probability. *Journal of Risk and Uncertainty, 8*, 167–196.
Carbone, E., & Hey, J. (1995). A comparison of the estimates of EU and non-EU preference functionals using data from pairwise choice and complete ranking experiments. *Geneva Papers on Risk and Insurance Theory, 20*, 111–133.
Delquié, P., & Cillo, A. (2006). Disappointment without prior expectation: A unifying perspective on decision under risk. *Journal of Risk and Uncertainty, 33*, 197–215.
Diecidue, E., & Wakker, P. P. (2001). On the intuition of rank-dependent utility. *Journal of Risk and Uncertainty, 23*, 281–298.
Ebert, S. (2015). On skewed risks in economic models and experiments. *Journal of Economics Behavior and Organization, 112*, 85–97.
Fechner, G. (1860). *Elements of psychophysics*. NewYork: Holt, Rinehart and Winston.
Fishburn, P. (1978). A probabilistic expected utility theory of risky binary choices. *International Economic Review, 19*, 633–646.
Gini, C. (1912), *Variabilità e Mutabilità, contributo allo studio delle distribuzioni e delle relazione statistiche: 3* (pp. 3–159). Studi economici-giuridici dela Regia Università di Cagliari.
Goldstein, W. M., & Einhorn, H. J. (1987). Expression theory and the preference reversal phenomena. *Psychological Review, 94*, 236–254.
Hey, J., & Carbone, E. (1995). Stochastic choice with deterministic preferences: An experimental investigation. *Economics Letters, 47*, 161–167.
Hey, J. D., & Orme, C. (1994). Investigating generalisations of expected utility theory using experimental data. *Econometrica, 62*, 1291–1326.
Hey, J. (1995). Experimental investigations of errors in decision making under risk. *European Economic Review, 39*, 633–640.
Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society, Series B, 52*, 105–124.
Köbberling, V., & Wakker, P. P. (2003). Preference foundations for nonexpected utility: A generalized and simplified technique. *Mathematics of Operations Research, 28*, 395–423.
Köbberling, V., & Wakker, P. (2005). An index of loss aversion. *Journal of Economic Theory, 122*, 263–291.
Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, XLVII*, 263–291.
Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of Measurement, Vol. 1 (Additive and Polynomial Representations)*. NY: Academic Press.
Loomes, G., & Sugden, R. (1998). Testing different stochastic specifications of risky choice. *Economica, 65*, 581–598.
Luce, R. D. (1996). The ongoing dialog between empirical science and measurement theory. *Journal of Mathematical Psychology, 40*, 78–98.
Machina, M. (2009). Risk, ambiguity, and the rank-dependence axioms. *American Economic Review, 99*(1), 385–392.
Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance, 7*(1), 77–91.
Prelec, D. (1998). The probability weighting function. *Econometrica, 66*, 497–527.
Quiggin, J. (1981). Risk perception and risk aversion among Australian farmers. *Australian Journal of Agricultural Recourse Economics, 25*, 160–169.
Safra, Z., & Segal, U. (1998). Constant risk aversion. *Journal of Economic Theory, 83*, 19–42.
Schmidt, U., Starmer, C., & Sugden, R. (2008). Third-generation prospect theory. *Journal of Risk and Uncertainty, 36*, 203–223.

Shalit, H., & Yitzhaki, S. (1984). Mean-Gini, portfolio theory, and the pricing of risky assets. *The Journal of Finance, 39*(5), 1449–1468.
Starmer, Ch. (2000). "Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk". *Journal of Economic Literature, 38*, 332–382.
Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*, 297–323.
Wakker, P., & Deneffe, D. (1996). Eliciting von Neumann–Morgenstern utilities when probabilities are distorted or unknown. *Management Science, 42*, 1131–1150.
Wakker, P. (1988). The algebraic versus the topological approach to additive representations. *Journal of Mathematical Psychology, 32*, 421–435.
Wilcox, N. (2008). Stochastic models for binary discrete choice under risk: A critical primer and econometric comparison. In J. C. Cox, & G. W. Harrison (Eds.), *Risk aversion in experiments: Vol. 12, Research in experimental economics* (pp. 197–292). Bingley, UK: Emerald.
Wilcox, N. (2010). 'Stochastically more risk averse:' A contextual theory of stochastic discrete choice under risk. *Journal of Econometrics, 162*, 89–104.
Yaari, M. (1987). The dual theory of choice under risk. *Econometrica, 55*, 95–115.
Yitzhaki, S. (1982). Stochastic dominance, mean variance, and Gini's mean difference. *The American Economic Review, 72*(1), 178–185.