Ordering higher risks in Yaari's dual theory

A. Castaño-Martínez, G. Pigueiras, C.D. Ramos, M.A. Sordo

Departamento de Estadística e Investigación Operativa
Facultad de Ciencias, Universidad de Cádiz

E-mail addresses: antonia.castano@uca.es, gema.pigueiras@uca.es,
carmen.ramos@uca.es mangel.sordo@uca.es

This version: September 3, 2022

Abstract

In Yaari's (1987) dual theory of choice under risk, risks preferences are based on a functional that depends on a subjective function called distortion. In the context of Wang's (1996) premium principle, Wang and Young (1998) considered a sequence of classes of partial orderings of risk distributions characterizing the preferences of groups of risk averse agents that base decisions on this functional. Under this approach, if a distribution is perceived as less risky than another, the mean of the former is smaller or equal to the latter's, making some risk distributions of interest non-comparable. In this paper, we study a sequence of partial orders of risk distributions based on comparisons of successive integrals of TVaR curves that characterize the preferences of groups of agents exclusively concerned with large risks higher than the expected values.

JEL: G22

Keywords: risk distribution, stochastic dominance, dual theory, distortion function

1 Introduction

The dual theory of choice under risk proposed by Yaari (1987) is an alternative to the classical expected utility theory obtained when the von Neumann and Morgenstern (1944) independence axiom is replaced by a dual axiom that induces a change in the structure of decision. Whereas expected utility assigns a value to a prospect (gain or loss) by using a transformed expectation that is linear in probabilities but nonlinear in wealth, the dual theory takes a transformed expectation that is linear in wealth but nonlinear in probabilities. In this framework, attitudes towards risk are characterized by a function (called distortion) that modifies the underlying probabilities before calculating a transform (or distorted) expectation, which is the certainty equivalent to the risk.

In actuarial, the dual theory has been applied to construct premium functionals for insurance contracts (Denneberg 1990, Wang 1996). For concave distortions, under suitable assumptions, these premiums (or Yaari functionals) have a representation as mixtures of TVaRs (tail values at risk), which are risk measures widely used in insurance. This representation suggests exploiting the TVaR curve (which is the curve defined by the tail value at risk at probability p for all probabilities) for comparing risks in the dual framework. Wang and Young (1998) show that non-intersecting TVaR curves induce a partial ordering of risk distributions equivalent to the order based on Yaari functionals with concave distortions. When the TVaR curves intersect, Wang and Young (1998) restrict attention to a nested family of appropriate concave distortions giving rise to a sequence of successively weaker orderings of risks. However, under Wang and Young's approach, if a distribution is perceived as less risky than another, the mean of the former is smaller or equal to the latter's, making some risk distributions of interest non-comparable. This limitation can also be explained in terms of risk preferences: the approach is inadequate to compare the shared preferences of strong risk averse agents who always use premiums strictly higher than the net premium to value their risks. To overcome this limitation, we consider a sequence of partial orderings of risks based on successive integrals of TVaR curves and provide an interpretation of these orderings in terms of Yaari functionals with distortions that describe higher degrees of risk aversion, compared to those based on any concave distortion. This task requires considering the dual counterpart of some concepts of risk aversion developed so far within the expected utility model, a topic that is interesting in its own right.

2 Motivation and previous notions

Let X be a risk (or non-negative random variable) with distribution function $F(x)$ and tail function $\bar{F}(x) = 1 - F(x)$. In Yaari's framework an agent uses a distortion g (that is, a non-decreasing function g from $[0, 1]$ to $[0, 1]$ such that $g(0) = 0$ and $g(1) = 1$) that transforms $\bar{F}(x)$ into $g(\bar{F}(x))$ and values the risk X at its certainty equivalent or distorted expectation, given by

$$H_g(X) = \int_0^\infty g(\bar{F}(x))dx. \tag{1}$$

In the context of the premium principle of Wang (1996), $H_g(X)$ represents the market price for transferring the risk X, that is, the sum of money which, when received with certainty, is considered by the agent equally as good as the risk X. In this setting, Wang (1996) shows that g is concave when the agent is risk averse. In such a case, $H_g(X) \geq E(X)$ and (1) is a coherent risk measure in the sense of Artzner et al. (1999). Moreover, it follows from Yaari axioms[1] that an agent with distortion g prefers X to Y (or is indifferent between them) if, and only if, $H_g(X) \leq H_g(Y)$, which reflects preference for less risky distributions.

An example of distorted expectation with concave distortion is the tail value at risk (or TVaR), defined by

$$\text{TVaR}_p(X) = \frac{1}{1-p}\int_p^1 F^{-1}(t)dt, \quad p \in [0,1),$$

where $F^{-1}(p) = \inf\{x : F(x) \geq p\}$, $0 \leq p \leq 1$, obtained by taking $g(t) = \min\{\frac{t}{1-p}, 1\}$ in (1). The tail value at risk induces the following partial order of risks: given two risks X and Y, we say that X is smaller than Y in the stop-loss order (denoted $X \leq_{sl} Y$) if $\text{TVaR}_p(X) \leq \text{TVaR}_p(Y)$ for all $p \in [0,1)$. The stop-loss order characterizes the behavior of agents that use Yaari functionals with concave distortions (Wang and Young, 1998).

Observe that if $X \leq_{sl} Y$, then necessarily $E[X] \leq E[Y]$, where $E[X]$ and $E[Y]$ denote the respective expectations of X and Y. As a consequence, the stop-loss order does not always capture the preferences of those agents who are more concerned about the extreme tail behavior of the risk distributions than mean values. There are many examples where X is considered less risky than Y when comparing their means and more risky when comparing their right-tails. To give one example, let X and Y be two Pareto risks with parameters $\alpha_1 = 2.4, \beta_1 = 3$ and $\alpha_2 = 2, \beta_2 = 2$, respectively (where the tail function of a Pareto random variable with parameters (α, β) is $\bar{F}(x) = \beta^\alpha (x+\beta)^{-\alpha}$ for $x > 0, \alpha > 1, \beta > 0$). Figure 1 plots the tail value at risk of X and Y as a function of p and shows that $\text{TVaR}_p(Y) \geq \text{TVaR}_p(X)$ for all $p > 0.8530$. For someone interested in avoiding large losses, the risk Y likely will be perceived as being more risky than X. The Yaari functional

$$H_1(X) = \int_0^\infty g_1(\bar{F}(x))dx, \tag{2}$$

[1] In the original Yaari setting, (1) is the valuation of a random economic prospect and g is convex when the agent is risk averse. In the context of the premium principle of Wang (1996), (1) is the valuation of a random loss and g is concave when the agent is risk averse, see Section 8.2 in Wang (1996).

with concave distortion

$$g_1(t) = \begin{cases} t(1-\log t), & \text{for } 0 < t \leq 1 \\ 0, & \text{for } t = 0, \end{cases} \tag{3}$$

supports this perception ($H_1(X) = 5.8163 < 6 = H_1(Y)$). However, $X \nleq_{sl} Y$, because $E(X) = 2.143 > 2 = E(Y)$. It is natural, therefore, to investigate an order weaker than the order $\leq_{sl}$ to explain the behavior of the collective of individuals who perceives Y as being more risky than X.

Some orders weaker than the stop-loss order have been studied before. For example, Wang and Young (1998) considered a sequence of progressively more risk averse distortions to increase the number of comparable risks, giving rise to a sequence of stochastic orders[2] (denoted by $\leq_n, n = 1, 2, ...$) weaker than the stop-loss order for $n \geq 3$. However, $X \leq_n Y$ also implies $E(X) \leq E(Y)$ and this sequence of orders also does not help to explain the behavior of agents who perceive Y as being more risky than X in the above example.

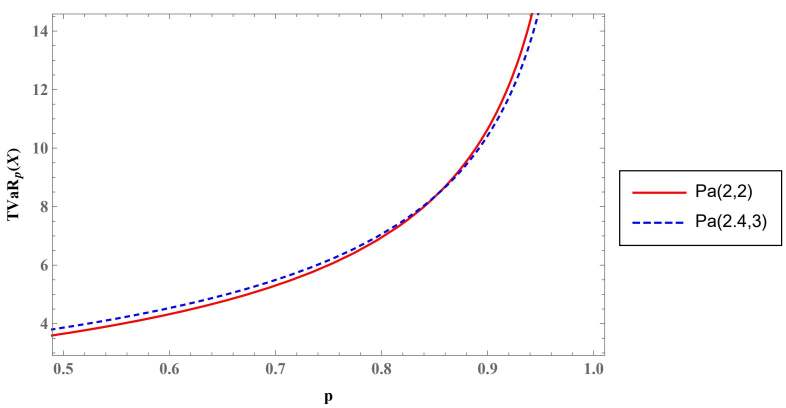

Figure 1: Tail value at risk of X and Y as a function of p

The choice of the distortion $g_1(t)$ in the above example is not arbitrary but motivated by its central role in this paper. Given a risk X, the Yaari functional $H_1(X)$ with distortion g_1 possesses appealing interpretations for actuaries. The first one comes from the representation

$$H_1(X) = \int_0^1 \text{TVaR}_p(X)dp, \tag{4}$$

(see Sordo et al., 2016), which offers a geometric interpretation of $H_1(X)$ as the area under the curve defined by $\text{TVaR}_p(X)$. Another interpretation is in the analysis of the time series of insurance claims. Given a sequence of independent claims having the same distribution as X, a claim X_j is an upper record claim if it is larger than all the previous claims of the sequence. In this context, $H_1(X)$ represents the first nontrivial expected upper record

[2]The formal definition of the order $\leq_n$ is given in Section 3 below.

claim (see Section 3.3 in Castaño-Martínez et al., 2020). It can also be shown (see Theorem 4 in Sordo et al., 2016) that

$$H_1(X) = E[X] + \epsilon(X), \tag{5}$$

where

$$\epsilon(X) = -\int_0^\infty \overline{F}(x) \log \overline{F}(x) dx$$

is a measure of variability called cumulative residual entropy (Rao et al., 2004) that has received growing interest among insurance researchers (see, for example, Sordo and Psarrakos 2017, Psarrakos and Sordo 2019, Hu and Chen 2020, Sun et al. 2022). In particular, the more variable the risk distribution, the further to the right from the mean is $H_1(X)$.

The rest of the paper is organized as follows. To overcome the drawback aforementioned, in this paper we define a sequence of stochastic orders, weaker than the stop-loss order, based on comparisons of successive integrals of the TVaR curve. The economic interpretation of these orders in the dual framework is given in terms of a nested sequence of classes $\{\widehat{D}_n, n \geq 3\}$ of distortions that are progressively more risk averse than the distortion g_1 given by (3), where the meaning of "g being progressively more risk averse than g_1" is explained in terms of coefficients of the form $g^{(n+1)}/g^{(n)}$ for $n = 1, 2, \ldots$. While the role of these coefficients as measures of higher-order risk attitudes in the utility framework (when g is a utility function) has received considerable attention (see, for example, Caballé and Pomansky 1996, Jindapon and Neilson 2007 and Denuit and Eeckhoudt 2010), their role in the dual framework, except for $n = 1$ (Yaari 1986, Eeckhoudt and Laeven 2021), has not been explored so far. Consequently, we face in Section 3 the interpretation of these coefficients within the dual theory. This task requires interpreting some notions and theorems taken from the expected utility framework in the dual setting. In light of the new dual concepts, we study and interpret in Section 4 the sequence $\{\widehat{D}_n, n \geq 3\}$ and provide several examples of risk measures of the form (1) with $g \in \widehat{D}_n$.

At this point, we recall that any risk measure of the form (1) associated with a concave distortion can be written as a mixture of TVaRs (Rockafellar et al. 2006, Pflug and Römisch 2007) as follows[3]

$$I_h(X) = \int_0^1 \text{TVaR}_p(X) dh(p), \tag{6}$$

with h being a weight function[4] (that is, a non-decreasing function from $[0, 1]$ to $[0, 1]$ with $h(0) = 0$ and $h(1) = 1$). Using this representation, we show in Section 5 that there is a one-to-one correspondence between the sequence of orderings by the classes $\{\widehat{D}_n, n \geq 3\}$ and a sequence of orderings based

[3]See also Sordo et al. (2016) and Castaño-Martínez et al. (2019).

[4]Although h is formally defined as a distortion function, we prefer not to call it distortion in this context, because h is not directly distorting $\bar{F}$.

on evaluation measures of the form (6) by certain classes of weight functions. We use this one-to-one correspondence in Section 6, where we characterize the risk attitudes of agents with distortions in $\widehat{D}_n$ by the announced distortion-free ordering of distributions based on successive integrals of TVaR curves. Conclusions are in Section 7.

Throughout the paper, given a function f, $f^{(k)}$ denotes the kth derivative of f, $k = 1, 2, \ldots,$ $f^{(0)} = f$. Sometimes, we write $f', f'', \ldots$ instead of $f^{(1)}, f^{(2)}, \ldots$. The notation $f(t) \uparrow$ means that $f(t)$ is non-decreasing in t.

3 Some notions of risk aversion in the dual framework

In this section, we reinterpret in our dual framework some notions and theorems[5] about risk aversion taken from the expected utility setting. Recall that, in the context of the premium principle of Wang (1995, 1996), the Yaari functional H_g defined by (1) is the valuation of a random loss, and an agent with distortion g prefers X to Y (or is indifferent between them) if $H_g(X) \leq H_g(Y)$.

First, we translate to our context the two concepts of one random variable Y having more nth degree risk than random variable X and a person being nth degree risk averse. Given a random variable X with distribution function F, we denote $F_1^{-1}(p) = F^{-1}(p)$ and

$$F_{n+1}^{-1}(p) = \int_p^1 F_n^{-1}(t)\,dt, \text{ for } n = 1, 2, \ldots, \ 0 \leq p \leq 1.$$

The following notion is the dual counterpart of a concept introduced by Ekern (1980) in the expected utility framework.

Definition 1 *Given two risks X and Y with respective distribution functions F and G and $n \geq 1$, we say that Y has more nth degree dual risk than X if*
(a) $F_k^{-1}(0) = G_k^{-1}(0)$, for $k = 1, \ldots, n$.
(b) $F_n^{-1}(p) \leq G_n^{-1}(p)$ for all $p \in [0, 1]$, with strict inequality for some p.

To interpret this definition, note that $F_k^{-1}(0) = \frac{1}{(k-1)!} E[\max(X_1, \ldots, X_{k-1})]$, where $k \geq 2$ and $X_1, \ldots, X_{k-1}$ are independent copies of X. Similarly, $G_k^{-1}(0) = \frac{1}{(k-1)!} E[\max(Y_1, \ldots, Y_{k-1})]$, where $Y_1, \ldots, Y_{k-1}$ are independent copies of Y. As in Wang and Young (1998), we refer to these expectations as the kth "dual moments" of X and Y, respectively[6]. When (a) in Definition 1 is satisfied, the $(n-1)$ first dual moments of X and Y are equal. Condition (b) implies $E[\max(X_1, \ldots, X_n)] < E[\max(Y_1, \ldots, Y_n)]$.

In words, for $n = 1$, Definition 1 says that Y has more first degree dual risk than X if and only if X is smaller than Y in the usual stochastic order

[5] The proofs of the results in this section follow the same pattern as those in the expected utility framework with convenient modifications. They have been included in Appendix A for the sake of completeness.

[6] In insurance, the kth dual moment of X can be thought as the expected higher claim in a set of k independent claims with the same distribution as X.

(which implies, in particular, that the mean of X is less than or equal to the mean of Y). For $n = 2$, Definition 1 says that Y has more second degree dual risk than X if and only if X is smaller than Y in the convex order[7] (denoted by $X \leq_{cx} Y$). The latter can be equivalently rewritten (see Theorem 3.A.1 in Shaked and Shanthikumar, 2007) as

$$\int_{-\infty}^{x} F(u)\mathrm{d}u \leq \int_{-\infty}^{x} G(u)\mathrm{d}u, \quad \text{for all } x,$$

which is equivalent to saying that G can be obtained from F by a sequence of one or more mean-preserving spreads in the sense of Rothschild and Stiglitz (1970).

Roughly speaking, for $n \geq 2$, the idea of Definition 1 is as follows: Y has more nth degree dual risk than X if Y is more variable (about the same mean) than X in a stochastic sense that depends on n in such a way that the extreme values in the right-tails of their distributions are given more weight as n increases. Let denote $GI_n(X) = E[\max(X_1, ..., X_n)] - E[X]$. In general, the assumption that Y has more nth degree dual risk than X implies that $E[X] = E[Y], GI_2(X) = GI_2(Y), ..., GI_{n-1}(X) = GI_{n-1}(Y)$ and $GI_n(X) < GI_n(Y)$, where $GI_n(X)$ is an extended Gini variability index that gives more weight to the extreme right-tails than the usual Gini mean difference (given by GI_2).

Many papers in the expected utility framework have been devoted to interpreting the signs of successive derivatives of the utility functions, including the key papers by Menezes et al. (1980), Ekern (1980), Kimball (1990) and Eeckhoudt and Schlesinger (2006). Although less attention has been paid to the interpretation of successive derivatives of distortions in the dual context, some research has also been conducted (including the papers by Muliere and Scarsini 1989, Wang and Young 1998, Chateauneuf et al. 2002 and Eeckhoudt et al. 2020). Some results in Muliere and Scarsini (1989) suggest a link between the sign of nth derivative of the distortion and the stochastic order given by Definition 1. In order to formalize this fact, let g be a distortion differentiable at least n times, with $n \geq 1$. The following definition is the dual analogous in our setting of the nth degree risk aversion notion considered in the expected utility framework by Ekern (1980).

Definition 2 *An agent with a distortion g is nth degree dual risk averse*[8] *if g is at least n times differentiable, with $n \geq 1$ and $(-1)^{n+1}g^{(n)}(x) \geq 0$, for all $x \in (0, 1)$.*

Observe that risk aversion in the sense of a concave distortion is indicated by $n = 2$ in the above definition. Note also that an agent with the distortion

[7] X is smaller than Y in the convex order if $E[X] = E[Y]$ and $X \leq_{sl} Y$ (see Section 3.4 in Denuit et al., 2005)

[8] We sometimes use the phrase "distortion g is nth degree dual risk averse" instead of the longer phrase "an agent with a distortion g is nth degree dual risk averse".

$g_1(t)$ defined by (3) exhibits nth degree dual risk aversion for any positive integer n.

The following result is the analogous counterpart of a result given by Ekern (1980). It shows the equivalence of Y having more nth degree dual risk than X and of X being preferred by every nth degree dual risk averter. The proof is in Appendix A.

Theorem 3 *For $n \geq 1$, Y has more nth degree dual risk than X if and only if every nth degree dual risk averter prefers X to Y.*

For $n = 2$, $n = 3$ and $n \geq 4$, respectively, Theorem 3 is the equivalent in our dual framework to the results given by Rothschild and Stiglitz (1970), Menezes et al. (1980) and Ekern (1980) in the expected utility context. For other interpretations of the sign of the kth derivative of a distortion, see Eeckhoudt et al. (2020). These authors provide an interpretation in terms of preferences between classes of lottery pairs with equal $k-1$ dual moments.

Wang and Young (1998) consider a sequence of progressively more risk averse classes of distortions to increase the number of comparable risks.

Definition 4 *Let $D_n, n \geq 1$, be the class of agents with a distortion g such that g is kth degree dual risk averse for $k = 1, 2, ..., n-1$ and $(-1)^n g^{(n-1)}$ is non-increasing*[9].

An example of distortion $g \in D_n$, for $n = 1, 2, ...$ is the distortion g_1 given by (3). Observe that $D_{n+1} \subset D_n$ for $n = 1, 2, ...$ According to Definition 2, an agent with a distortion $g \in D_n$ reaches higher degrees of risk aversion as n increases[10]. The risk attitudes of agents in D_n can be characterized by the following distortion-free ordering of distributions.

Definition 5 *Given two risks X and Y with distribution functions F and G, respectively, X is said to be smaller than Y in the nth dual stochastic order (denoted by $X \leq_n Y$) if*
(a) $F_k^{-1}(0) \leq G_k^{-1}(0)$, for $k = 1, ..., n$.
(b) $F_n^{-1}(p) \leq G_n^{-1}(p)$ for all $p \in [0, 1]$.

The nth dual stochastic order was first studied by Muliere and Scarsini (1989). For $n = 1, 2$, it reduces to the usual stochastic order and the stop-loss order, respectively. The following result is due to Wang and Young (1998) (see Theorem 4.4).

Theorem 6 *$X \leq_n Y$ if and only if $H_g(X) \leq H_g(Y)$, for all $g \in D_n$.*

[9] Wang and Young (1998) also include in $D_n, n \geq 2$ the distortion function $g_p(t) = 1 - (\frac{p-t}{p})_+^{n-1}, 0 < p \leq 1$, which has piecewise continuous $(n-1)$th derivative.

[10] Wang and Young (1998) provide their own interpretation about how the degree of risk aversion of $g \in D_n$ becomes higher as n increases.

Theorem 6 provides motivation, both from the economic and statistical approaches, for the sequence of the distortion-free partial ordering of risks given in Definition 5. However, $X \leq_n Y$ implies $E(X) \leq E(Y)$, for $n = 1, 2, ...$ Therefore, the nth dual stochastic order suffers, for $n \geq 3$, from the same shortcoming as the stop-loss order when the interest is in capturing the preferences of those agents that base their decisions on the right-tail risk rather than on the mean of a distribution. To overcome this shortcoming, in the next section, we will focus on "the most risk averse" agents in D_n, for $n \geq 3$. Let us first introduce a dual notion of risk aversion to explain what we mean when we say "the most risk averse" agents in D_n.

Whereas the signs of successive derivatives of a distortion inform us about risk preferences, they do not give any information about their intensity. The strength of risk aversion in the expected utility model is locally evaluated in terms of a quotient of derivatives of the utility function. In the dual context, Yaari (1986) has proposed to measure risk aversion by the quotient g''/g' (where g is a twice differentiable strictly increasing distortion), which is the analog of the Arrow-Pratt index of absolute risk aversion used in the expected utility model, see Pratt (1964) and Arrow (1970). This coefficient can be generalized as follows.

Definition 7 *Let g be a distortion $n+1$ times differentiable, $n \geq 1$. We define the nth degree coefficient of dual risk aversion as*

$$r_{g,n}(t) = \frac{-g^{(n+1)}(t)}{g^{(n)}(t)}, \quad t \in (0,1). \tag{7}$$

This coefficient can be used to define a distortion h as being more nth degree dual risk averse than a distortion g.

Definition 8 *Given two distortions h and g, we say that h is more nth degree dual risk averse than g if $r_{h,n}(t) \geq r_{g,n}(t)$ for all $t \in (0,1)$.*

The coefficient $r_{g,1}$ is the dual counterpart of the classical Arrow- Pratt index of absolute risk aversion, whereas $r_{g,2}$ is the dual of the Kimball index of prudence. For $n \geq 3$, $r_{g,n}$ is the dual version of the index of absolute risk aversion introduced by Caballé and Pomansky (1996) and studied, in the expected utility model, by Chiu (2005), Jindapon and Neilson (2007), Denuit and Eeckhoudt (2010) and Wei (2017). To interpret this index in the dual context, we follow the same comparative statics approach used in the expected utility model by Jindapon and Neilson (2007). We need the following definition.

Definition 9 *We say that Y differs from X by a simple increase in nth degree dual risk if*
(a) Y has more nth degree dual risk than X.
(b) There exists $p_0 \in [0,1]$ such that $G^{-1}_{n-1}(p) \leq F^{-1}_{n-1}(p)$ for all $p \leq p_0$ and $G^{-1}_{n-1}(p) \geq F^{-1}_{n-1}(p)$ for all $p \geq p_0$.

Observe that the property of Y differing from X by a simple increase in nth degree dual risk is a bit stronger than the property of Y having more nth degree dual risk than X (condition (b) in Definition 9, together with $F_n^{-1}(0) \leq G_n^{-1}(0)$, implies condition ($b$) in Definition 1). This property can be interpreted in terms of (dual) precedence relations as in Chiu (2005). We can now address the interpretation of the local coefficient $r_{g,n}(t)$ given by (7). Consider two Yaari risk averse agents with distortions i and j facing a risk Y. Let X be another risk which both agents strictly prefer to Y, with X and Y being comonotonic[11]. From Yaari's dual independence axiom it follows that, for any $t \in (0, 1]$, the agents prefer the risk $Z_t = tX + (1-t)Y$ to the risk Y and that Z_{t_1} is preferred to Z_{t_2} whenever $t_1 > t_2$. Now suppose that the agents can improve their risk positions from Y to Z_t for a cost of $c(t)$, where c is assumed to be increasing and convex, with $c(0) = 0$ and $c(1) = 1$. If $\bar{M}_t$ is the tail function of Z_t, it seems reasonable that the first and the second agents choose t to minimize

$$U_i(t) = \int_0^1 i(\bar{M}_t(p))dp + c(t) \tag{8}$$

and

$$U_j(t) = \int_0^1 j(\bar{M}_t(p))dp + c(t), \tag{9}$$

respectively.

The distortions i and j are assumed to be strictly increasing, concave and twice continuously differentiable. We also assume that $U_i'' \geq 0$ and $U_j'' \geq 0$ so that the first-order conditions identify the optimal values t_i and t_j, respectively. Intuition suggests that if the agent with distortion j is more risk averse than the agent with distortion i, then $t_j \geq t_i$ (which means that the first one is willing to spend more to improve the distribution[12]).

The following result shows that the more nth degree dual risk averse agent chooses a less risky but more costly distribution. The proof follows the same steps as the proof of Theorem 3 in Jindapon and Neilson (2007). We include it in Appendix A.

Theorem 10 *Consider two agents nth and $(n-1)$th degree dual risk averse, for $n \geq 2$, with distortions i and j strictly increasing and infinitely continuously differentiable. Let t_i and t_j minimize* (8) *and* (9)*, respectively. For any risk Y which differs from X by a simple increase in nth degree dual risk, $t_i \leq t_j$ if and only if j is more nth degree dual risk averse than i.*

4 Some classes of risk averse agents

We introduce in this section a sequence of classes $\{\widehat{D}_n, n \geq 3\}$ of risk averse agents (or risk averse distortions). By construction, $\widehat{D}_n \subset D_n$, where $\{D_n, n \geq$

[11] X and Y are comonotonic if there exists a risk W and two non-decreasing functions f and g such that $X = f(W)$ and $Y = g(W)$.

[12] This comparative problem is the adaptation to our dual context of the original problem considered in the expected utility framework by Jindapon and Neilson (2007).

1} is the sequence considered by Wang and Young (1998) (see Definition 4 above). The distortion $g_1(t)$ given by (3) plays a key role in the construction of the new sequence. Since $g_1(t) \in D_n$ for all positive integer n, the idea is that an agent with distortion g in D_n belongs to $\widehat{D}_n$ if they are more risk averse than an agent with the distortion $g_1(t)$, where risk aversion is measured by the coefficient of dual risk aversion (7).

Definition 11 *For $n \geq 3$, let $\widehat{D}_n$ be the class of distortions g such that g is at least $n-1$ times differentiable, $g'(1) = 0$,*

$$\frac{-tg^{(k+1)}(t)}{g^{(k)}(t)} \geq k-1, \quad t \in (0,1), \;\; k = 1, ..., n-2, \tag{10}$$

and $(-1)^n(tg^{(n-1)}(t) + (n-3)g^{(n-2)}(t))$ is non-increasing in t.

It can be checked that $\widehat{D}_n \subset D_n$ for $n \geq 3$. To interpret the succession $\{\widehat{D}_n, n \geq 3\}$ we first focus on $\widehat{D}_3$, the class of distortions g such that g is at least 2 times differentiable and such that $g'(1) = 0$, $g''(t) \leq 0$ for all $t \in (0,1)$ and $tg''(t) \uparrow$. An agent using a distortion in $\widehat{D}_3$ overweights large losses (in the sense that the relative loading $g(\overline{F}(x))/\overline{F}(x)$ increases faster as x goes to 0 if $g'(1) = 0$). Since the identity function does not belong to $\widehat{D}_3$, an agent with a distortion in $\widehat{D}_3$ does not use the mean value as a basis for evaluating risks. To interpret the meaning of $tg''(t) \uparrow$, assume that the third derivative of g exists and observe that $r_{g_1,2}(t) = \frac{1}{t}$, where $g_1(t)$ is the distortion given by (3). Then, $tg''(t) \uparrow$ if and only if $r_{g,2}(t) \geq r_{g_1,2}(t)$ for all $t \in (0,1)$. This means, according to Definition 8, that g is more 2nd degree dual risk averse than g_1. Since $\widehat{D}_3 \subset D_3$ (where D_3 is the class of concave distortions that also have increasing second derivatives, see Definition 4 above), we can think that agents with distortions in $\widehat{D}_3$ are among the most risk averse agents with distortions in D_3, in the sense that they do not use mean values for evaluating risks, overweight large losses and are more 2nd degree dual risk averse with respect to g_1 than other agents that are not in the same class.

For $n \geq 4$, condition (10) means that g is more kth degree dual risk averse than g_1 for $k = 1, ..., n-2$. Moreover, if g is n times differentiable, the last condition in the statement means that g is more $(n-1)$th degree dual risk averse than g_1. Since the classes are nested (that is, $\widehat{D}_{n+1} \subset \widehat{D}_n$ for $n \geq 3$) this implies that an agent with a distortion $g \in \widehat{D}_{n+1}$ is more dual risk averse with respect to g_1 than an agent with a distortion in $\widehat{D}_n$ (but not in $\widehat{D}_{n+1}$) in the sense of Definition 8. Since g_1 belongs trivially to $\widehat{D}_n$ for $n \geq 4$, we can think that agents with distortions in $\widehat{D}_n$ are among the most risk averse agents with distortions in D_n, in the sense that they do not use mean values for evaluating risks, overweight large losses and are more kth degree dual risk averse with respect to g_1, for $k = 1, ..., n-2$, than other agents that are not in the same class. Obviously, for agents with distortions in $\widehat{D}_n$, right-tail risks matter more than small risks.

Next, we provide several examples of risk measures of the form $H_g(X)$ with $g \in \widehat{D}_n$.

4.1 The class $H_m(X), m \geq 1$

For $m \geq 1$ the distortion

$$g_m(t) = mt \int_t^1 \frac{(1-u)^{m-1}}{u} du + 1 - (1-t)^m, \ 0 < t \leq 1, \tag{11}$$

belongs to $\widehat{D}_n$ for $n \geq 3$. The corresponding distortion risk measure (Sordo et al., 2016) is[13]

$$\begin{aligned} H_m(X) &= \int_0^1 g_m(\overline{F}(t))dt \\ &= m \left\{ H_1(X) - \sum_{j=2}^{m} \frac{E[\max(X_1, ..., X_j)]}{j(j-1)} \right\} \\ &= E[X|X > max\{X_1, ..., X_m\}], \ m \geq 2, \end{aligned} \tag{12}$$

where $H_1(X)$ is given by (2) and $X_1, ..., X_m$ are independent copies of X. For $m \geq 1$ and $0 < t \leq 1$ it is shown in Sordo et al. (2016) that

$$g_{m+1}(t) = g_m(t) + t \sum_{j=m+1}^{\infty} \frac{(1-t)^j}{j},$$

which implies that $H_m(X) \leq H_{m+1}(X)$ (an insurer with risk averison to higher risks may use $H_{m+1}(X)$ instead of $H_m(X)$ to evaluate X).

4.2 Record claims

Given $j \geq 1$, the distortion

$$g_j(t) = t \sum_{k=0}^{j} \frac{(-log\, t)^k}{k!}, \quad 0 < t \leq 1, \tag{13}$$

where $g_j(0) = 0$ is defined by continuity. In particular, for $j = 1$, (13) reduces to (3). It can be checked that $g_j \in \widehat{D}_n$ for all $n \geq 3$ and $j \geq 1$. The corresponding distortion risk measure can be interpreted in terms of the upper record claims[14] $\{R_j(X)\}_{j\geq 1}$ of a sequence of independent claims, $\{X_i\}_{i\geq 1}$, having the same distribution as X. Specifically,

$$H_{g_j}(X) = E[R_{j+1}(X)]$$

is the expected $(j+1)$th record claim, with $j \geq 1$. In particular, the risk measure (4) can be interpreted as the expected first nontrivial record claim (for details, see Section 3.3 in Castaño-Martínez et al., 2020).

[13]More formally written, $H_m(X) = E[Y_X]$ where

$$[Y_X \mid max\{X_1, ..., X_m\} = x] \stackrel{d}{=} [X \mid X > x] \quad \text{for all } x \geq 0.$$

[14]Given a sequence of independent claims having the same distribution as X, we say that X_j is an upper record claim if it is larger than all the previous claims of the sequence.

4.3 Modified PH-transform risk measure

Let $m \geq 2$. The distortion

$$g_m(t) = \frac{mt^{1/m} - t}{m-1}, \quad 0 \leq t \leq 1, \tag{14}$$

belongs to $\widehat{D}_n$, for $n \geq 3$. It is not difficult to see from (14) that $g_m(t) \leq g_{m+1}(t)$ for $m \geq 2$, which implies that $H_{g_{m+1}}(X)$ gives a higher premium than $H_{g_m}(X)$. One of the alternative representations for the distortion risk measure is

$$H_{g_m}(X) = \int_0^\infty \left(\bar{F}(x)\right)^{1/m} dx + \frac{1}{m-1}\int_0^\infty \left(\left(\bar{F}(x)\right)^{1/m} - \bar{F}(x)\right) dx,$$

which decomposes $H_{g_m}(X)$ as a sum of the PH-transform risk measure (Wang, 1995) plus a deviation measure from the mean

$$\frac{1}{m-1}\int_0^\infty \left(\left(\bar{F}(x)\right)^{1/m} - \bar{F}(x)\right) dx$$

(if $m = 2$, this is the right-tail deviation proposed by Wang 1998). For $m \geq 2$, $H_{g_m}(X)$ has a simple analytical form in several parametric families of distributions:

a) If X is exponentially distributed with mean λ then $H_{g_m}(X) = \lambda(1+m)$.

b) If X is uniformly distributed over $(0, a)$, then $H_{g_m}(X) = \frac{a}{2}\left[1 + \frac{m}{m+1}\right]$.

c) If X is a Pareto random variable with parameters (α, β), then $H_{g_m}(X) = \frac{\beta}{\alpha-1}\left[1 + \frac{m\alpha}{\alpha-m}\right]$, for $\alpha > m$.

5 Two sequences of ordering of risks

The purpose of this section is to show that a risk ordering based on Yaari functionals of the form (1) with distortions in $\widehat{D}_n$ is equivalent to a risk ordering based on mixtures of TVaRs of the form (6) with mixing weight functions that belong to a class C_n defined below. This result will be helpful in Section 6 to characterize the risk attitudes of agents with distortions in $\widehat{D}_n$ by a distortion-free ordering of distributions.

Definition 12 *Given $n \geq 1$, C_n is the class of weight functions h such that h is at least $n-1$ times differentiable, $h^{(k)} \geq 0$ for $k = 1, \ldots, n-1$ and $h^{(n-1)}$ is non-decreasing* [15].

[15] Given $0 \leq p < 1$, we also include in C_{n+1}, $n \geq 1$, the function

$$h_{p,n}(t) = \frac{(t-p)_+^n}{(1-p)^n}, \tag{15}$$

which has piecewise continuous n-th derivative.

It is well-known that a risk measure of the form (1) associated with a concave distortion can be written as a mixture of TVaRs of the form (6) (see[16] Rockafellar et al. 2006 and Pflug and Römisch 2007). Here we are interested in higher order relationships between the distortion g in (1) and the mixing weight function h in (6).

Theorem 13 *Let X and Y be two non-negative random variables and let $n \geq 2$. Then, $I_h(X) \leq I_h(Y)$ for all $h \in C_n$ if and only if $H_g(X) \leq H_g(Y)$ for all $g \in \widehat{D}_{n+1}$.*

Proof. The proof consists of showing that every functional of the form (6), with $h \in C_n, n \geq 2$, can be written in the form (1), with $g \in \widehat{D}_{n+1}$ and vice versa. For $n \geq 2$, every $I_h(X)$ with $h \in C_n$ can be written as

$$\begin{aligned} I_h(X) &= \int_0^1 \frac{1}{1-p} \int_p^1 F^{-1}(t) dt dh(p) \\ &= \int_0^1 \int_0^t \frac{1}{1-p} dh(p) F^{-1}(t) dt \\ &= \int_0^1 g_h'(1-t) F^{-1}(t) dt \\ &= \int_0^\infty g_h(\bar{F}(x)) dx, \end{aligned}$$

where

$$g_h'(t) = \int_0^{1-t} \frac{1}{1-p} dh(p). \tag{16}$$

Clearly $g_h'(1) = 0$. Since h is a weight function, $g_h'(t)$ is not increasing and $g_h'(t) \geq 0$ for all $t \in (0,1)$. Moreover, if we set $g_h(0) = 0$,

$$\begin{aligned} g_h(t) &= \int_0^t g_h'(u) du \\ &= \int_0^t \int_0^{1-u} \frac{1}{1-p} dh(p) du \\ &= \int_0^1 \int_0^{\min\{1-p,t\}} du \frac{1}{1-p} dh(p) \\ &= t \int_0^{1-t} \frac{1}{1-p} dh(p) + \int_{1-t}^1 dh(p) \\ &= 1 - h(1-t) + t \int_0^{1-t} \frac{1}{1-p} dh(p). \end{aligned} \tag{17}$$

[16] Rockafellar et al. (2006, Prop. 5) and Pflug and Römisch (2007, Prop. 2.64) study the conditions under which the functional $\int_0^1 F^{-1}(t) d\lambda(t)$ with λ concave can be written as a mixture of the form $\int_0^1 \mathrm{CVaR}_p(X) d\mu(p)$, where $\mathrm{CVaR}_p(X) = \frac{1}{p} \int_0^p F^{-1}(t) dt$ (and conversely).

It is easy to see that $g_h(t)$ is a distortion function. Now assume that $h \in C_n$, with $n \geq 2$. Hence, it can be shown inductively from (16) for $k = 2, ..., n$, that

$$g_h^{(k)}(t) = (-1)^{k+1}\left[\frac{h^{(k-1)}(1-t)}{t} + \frac{(k-2)h^{(k-2)}(1-t)}{t^2}\right.$$
$$\left. + \frac{(k-2)(k-3)h^{(k-3)}(1-t)}{t^3} + ... + \frac{(k-2)!h'(1-t)}{t^{(k-1)}}\right].$$

Since $h^{(i)}(1-t) \geq 0$ for $i = 1, ..., n-1$, it follows that $(-1)^{k+1}g_h^{(k)}(t) \geq 0$ for all $t \in (0,1]$ and $k = 1, \ldots, n$. Moreover, using that $h^{(n-i)}(1-t)/t^i$, $i = 1, ..., n-1$, is non-increasing, it follows that $(-1)^{n+1}g_h^{(n)}(t)$ is also non-increasing, which shows that $g_h \in D_{n+1}$. Now, it can be checked by taking successive derivatives of (16) that

$$h^{(k)}(t) = (-1)^k((1-t)g_h^{(k+1)}(1-t) + (k-1)g_h^{(k)}(1-t)), \quad k = 1, ..., n-1. \tag{18}$$

Since $h^{(k)}(t) \geq 0$ for all $t \in [0,1)$ and $k = 1, ..., n-1$, we see from (18) that $(-1)^k(tg_h^{(k+1)}(t) + (k-1)g_h^{(k)}(t)) \geq 0$ for all $t \in (0,1]$, $k = 1, ..., n-1$, which implies $tr_{g,k}(t) \geq k-1$ for $k = 1, ..., n-1$. Finally, since $h^{(n-1)}$ is non-decreasing, $(-1)^{n-1}((1-t)g_h^{(n)}(1-t) + (n-2)g_h^{(n-1)}(1-t))$ is non-decreasing, which means that $g_h \in \widehat{D}_{n+1}$.

To prove the converse, let $g \in \widehat{D}_n$ with $n \geq 3$. Then,

$$\begin{aligned} H_g(X) &= \int_0^\infty g(\bar{F}(x))dx \\ &= \int_0^1 g'(p)F^{-1}(1-p)dp \\ &= -\int_0^1 \left(\int_{1-p}^1 F^{-1}(t)dt\right) dg'(p) \\ &= \int_0^1 \mathrm{TVaR}_p(X)dh_g(p) \end{aligned}$$

where

$$h_g(p) = \int_0^p (1-t)dg'(1-t) \tag{19}$$

is a weight function. By taking successive derivatives in (19) we obtain

$$h_g^{(k)}(t) = (-1)^k((1-t)g^{(k+1)}(1-t) + (k-1)g^{(k)}(1-t)), \quad k = 1, ..., n-2. \tag{20}$$

Using sequentially in (20) that $g^{(k+1)}$ and $g^{(k)}$ have opposite signs and that $tr_{g,k}(t) \geq k-1$ for $k = 1, ..., n-2$, it follows that $h_g^{(k)}(t) \geq 0$ for all $t \in (0,1)$. Moreover, it follows from (20) and the assumptions on g that $h_g^{(n-2)}(t)$ is non-decreasing, therefore $h_g \in C_{n-1}$. ■

The following corollary shows that an agent with distortion in $\widehat{D}_n$ always uses a premium higher than $H_1(X)$. This, together with (5), suggests that this agent is concerned about the right-tail risk and not about losses around the mean value.

Corollary 14 *Let $g \in \widehat{D}_n, n \geq 3$. Then, $H_g(X) \geq H_1(X)$.*

Proof. Let $g \in \widehat{D}_n, n \geq 3$. From the proof of Theorem 13 we know that there exists $h_g \in C_2$ such that $H_g(X) = \int_0^1 \mathrm{TVaR}_p(X) dh_g(p)$. Since h_g is a convex weight function, $h_g(p) \leq p$ for all $p \in [0,1]$. Moreover, $\mathrm{TVaR}_p(X)$ is non-decreasing in p, therefore

$$\int_0^1 \mathrm{TVaR}_p(X) d(h_g(p) - p) = \int_0^1 (p - h_g(p)) d\mathrm{TVaR}_p(X) \geq 0,$$

that is, $H_g(X) \geq H_1(X)$. ■

6 A new stochastic dominance rule

This section aims to characterize the risk attitudes of agents with distortions in $\widehat{D}_n$ by a distortion-free ordering of distributions based on the following sequence of functions. Denote $\mathrm{T}_X^{[1]}(p) = \mathrm{TVaR}_p(X)$ and define

$$\mathrm{T}_X^{[n]}(p) = \int_p^1 \mathrm{T}_X^{[n-1]}(t) dt, \quad \text{for } n = 2, 3, \ldots \quad \text{and } 0 \leq p \leq 1.$$

By analogy with Definition 5, we define the following sequence of stochastic orders.

Definition 15 *Given X and Y two non-negative risks and $n \geq 2$, we write $X \leq_{tvar[n]} Y$ if*
(a) $T_X^{[k]}(0) \leq T_Y^{[k]}(0)$ for $k = 2, ..., n$.
(b) $T_X^{[n]}(p) \leq T_Y^{[n]}(p)$ for all $p \in [0,1]$.

To interpret Definition 15, we rewrite $\mathrm{T}_X^{[n]}(0)$ using a more conventional expression.

Lemma 16 *Let X be a random variable and $k \geq 2$. Then*

$$T_X^{[k]}(0) = \frac{1}{(k-1)!} E[X | X > max\{X_1, ..., X_{k-1}\}]$$

where $X_1, ..., X_{k-1}$ are independent copies of X.

Proof. By using induction on j, it is easy to prove that

$$\frac{j!}{(1-p)^j} \mathrm{T}_X^{[k]}(p) = \int_0^1 \mathrm{T}_X^{[k-j]}(t) dh_{p,j}(t), \quad j = 1, \ldots, k-1,$$

for all $p \in [0,1)$, where the functions $h_{p,k-1} \in C_k, k \geq 2$, are given in (15). By taking, in particular $j = k-1$, we obtain

$$\frac{(k-1)!}{(1-p)^{k-1}} T_X^{[k]}(p) = I_{h_{p,k-1}}(X), \tag{21}$$

where I_h is given by (6). Now consider the case $p = 0$. Since $h_{0,k-1}(t) = t^{k-1}$, it follows from (17) that $I_{h_{0,k-1}}(X) = H_{g_{k-1}}(X)$, where $g_{k-1}(t)$ is given by (11). The result follows from (12). ■

Thus, $\mathrm{T}_X^{[k]}(0)$ is, up to a scale factor, the distortion risk measure $H_{k-1}(X)$ considered in Section 4.1. Therefore, condition (a) in Definition 15 can be equivalent rewritten as

$$E[X|X > max\{X_1, ..., X_{k-1}\}] \leq E[Y|Y > max\{Y_1, ..., Y_{k-1}\}], \tag{22}$$

for $k = 2, ..., n$, where $X_1, ..., X_{k-1}$ are independent copies of X and $Y_1, ..., Y_{k-1}$ are independent copies of Y. Condition (b) in Definition 15 implies, in particular, that (22) holds for $k = n+1$. Consequently, an agent will interpret the order $\leq_{tvar[n]}$ as follows: if $X \leq_{tvar[n]} Y$, the variability of the risk along the right-tail distribution is smaller for X than for Y (and the higher n, the more the agent is focused on extreme losses).

The following expansion formula for $I_h(X)$ will be useful to prove the main result in this section.

Lemma 17 *Let $I_h(X)$ be a risk measure of the form* (6) *with $h \in C_n, n \geq 2$. Then,*

$$I_h(X) = \sum_{k=1}^{n-1} h^{(k)}(0) T_X^{[k+1]}(0) + \int_0^1 T_X^{[n]}(t) dh^{(n-1)}(t). \tag{23}$$

Proof. The result will be proved inductively on n using that $C_{n+1} \subset C_n$. Let $h \in C_2$, which means that h is differentiable, $h' \geq 0$ and h' is non-decreasing. Then,

$$\begin{aligned} I_h(X) &= \int_0^1 TVaR_t(X) dh(t) \\ &= -\int_0^1 h'(t) dT_X^{[2]}(t) \\ &= h'(0) T_X^{[2]}(0) + \int_0^1 T_X^{[2]}(t) dh'(t), \end{aligned}$$

where we have used integration by parts in the third equality together with the fact that $T_X^{[2]}(1) = 0$. Now suppose that (23) holds for all $h \in C_n$. Given $h \in C_{n+1}$, we must show that

$$I_h(X) = \sum_{k=1}^{n} h^{(k)}(0) T_X^{[k+1]}(0) + \int_0^1 T_X^{[n+1]}(t) dh^{(n)}(t). \tag{24}$$

Since $C_{n+1} \subset C_n$, h satisfies (23). Moreover, h is n times differentiable, so we can write

$$\int_0^1 T_X^{[n]}(t)dh^{(n-1)}(t) = -\int_0^1 h^{(n)}(t)dT_X^{[n+1]}(t)$$
$$= h^{(n)}(0)T_X^{[n+1]}(0) + \int_0^1 T_X^{[n+1]}(t)dh^{(n)}(t), \quad (25)$$

where we have used integration by parts again in the second equality together with the fact that $T_X^{[n+1]}(1) = 0$. Replacing (25) in (23) we obtain (24). ■

We can now prove the main result of this section.

Theorem 18 *Let X and Y be two non-negative random variables and let $n \geq 2$. Then, $X \leq_{tvar[n]} Y$ if only if $H_g(X) \leq H_g(Y)$ for all $g \in \widehat{D}_{n+1}$.*

Proof. By Theorem 13, we can prove equivalently that $X \leq_{tvar[n]} Y$ holds if and only if

$$I_h(X) \leq I_h(Y) \text{ for all } h \in C_n. \quad (26)$$

Let $I_h(X)$ and $I_h(Y)$ be two indices with $h \in C_n, n \geq 2$, and assume $X \leq_{tvar[n]} Y$. Then, $T_X^{[n]}(p) \leq T_Y^{[n]}(p)$ for all $p \in [0,1]$, $T_X^{[k]}(0) \leq T_Y^{[k]}(0)$ for $k = 2, ..., n$, $h^{(k)}(0) \geq 0$ for $k = 1, ..., n-1$ and $h^{(n-1)}$ is non-decreasing. Using (23), we see that these conditions imply (26), which proves the sufficiency.

To prove the converse, we assume that (26) holds for some $n \geq 2$. In particular, we have $I_{h_{p,n-1}}(X) \leq I_{h_{p,n-1}}(Y)$ for all $p \in [0,1)$ where $h_{p,n-1} \in C_n$ is given in (15). This, combined with (21), implies $T_X^{[n]}(p) \leq T_Y^{[n]}(p)$ for all $p \in [0,1]$. We also have from (26) that $I_{h_{0,k}}(X) \leq I_{h_{0,k}}(Y)$ for $k = 1, ..., n-1$, where $h_{0,k}(p) = p^k \in C_n$, for all $n \geq 1$. This, combined with (21), implies $T_X^{[k]}(0) \leq T_Y^{[k]}(0)$ for $k = 2, ..., n$ and $X \leq_{tvar[n]} Y$ holds. ■

Theorem 18 provides motivation, both from the economic and statistical approaches, for the sequence of distortion-free partial ordering of risks given in Definition 15: a risk X is smaller than Y in the order $\leq_{tvar[n]}$ if any agent with a distortion $g \in \widehat{D}_{n+1}$ perceives Y as being more risky than X.

The relationship between the nth dual stochastic order (Definition 5) and the ordering $\leq_{tvar[n]}$ follows immediately from Theorem 6 and Theorem 18 using that $\widehat{D}_n \subset D_n$ for $n \geq 2$.

Corollary 19 *Let X and Y be two non-negative random variables. For $n \geq 3$, $X \leq_n Y$ implies $X \leq_{tvar[n-1]} Y$.*

The condition $X \leq_{tvar[n]} Y$ can be weakened if we restrict to a subclass of distortions in $\widehat{D}_{n+1}$. Let us define

$$\widehat{D}_n^* = \{g \in \widehat{D}_n \text{ such that } g^{(k)}(1) = 0,\ k = 2, ..., n-1\},\ n \geq 3.$$

To cite some examples, the distortion g_m given by (11) belongs to $\widehat{D}_n^*$ whenever $m + 1 \geq n \geq 3$ and the distortion g_j given by (13) belongs to $\widehat{D}_n^*$ if $j \geq n-1$, with $n \geq 3$. If we restrict attention to the class $\widehat{D}_{n+1}^*$ the characterization in Theorem 18 can be stated as follows.

Corollary 20 *Let X and Y be two non-negative random variables and let $n \geq 2$. Then $T_X^{[n]}(p) \leq T_Y^{[n]}(p)$ for all $p \in [0,1]$ if and only if $H_g(X) \leq H_g(Y)$ for all $g \in \widehat{D}_{n+1}^*$.*

Proof. Define $C_n^* = \{h \in C_n \text{ such that } h^{(k)}(0) = 0,\ k = 1, ..., n-1\}, n \geq 2$. If follows directly from (18) and (20) that $g \in \widehat{D}_{n+1}^*$ if and only if $h_g \in C_n^*$, for $n = 2, 3, ...$ By following the same steps as in the proof of Theorem 18 and taking into account that $h^{(k)}(0) = 0,\ k = 1, ..., n-1$, for $h \in C_n^*$ we can prove that $T_X^{[n]}(p) \leq T_Y^{[n]}(p)$ for all $p \in [0,1]$ if and only if $I_h(X) \leq I_h(Y)$, which proves the result. ■

Of particular interest is the following corollary, which follows from Theorem 18 by taking $n = 2$.

Corollary 21 *Let X and Y be two non-negative random variables with distribution functions F_X and F_Y, respectively. Then,*

$$\int_p^1 TVaR_t(X)dt \leq \int_p^1 TVaR_t(Y)dt, \quad 0 \leq p \leq 1,$$

if and only if

$$\int_0^\infty g(\bar{F}_X(x))dx \leq \int_0^\infty g(\bar{F}_Y(x))dx,$$

for all distortions g such that $g'(1) = 0, g''(t) \leq 0, tg''(t) \uparrow$.

The following theorem shows that a "single-crossing property" on the curves $\mathrm{TVaR}_p(X)$ and $\mathrm{TVaR}_p(Y)$ implies the order $\leq_{tvar[2]}$.

Theorem 22 *Suppose $H_1(X) \leq H_1(Y)$ and there is some p_0 in $[0,1]$ such that $TVaR_p(X) \geq TVaR_p(Y)$ for p in $[0, p_0]$ and $TVaR_p(X) \leq TVaR_p(Y)$ for p in $[p_0, 1]$. Then $X \leq_{tvar[2]} Y$.*

Proof. By the assumptions,

$$\int_p^1 \mathrm{TVaR}_t(X)dt \leq \int_p^1 \mathrm{TVaR}_t(Y)dt \tag{27}$$

for all p in $[p_0, 1]$. If $\int_p^1 \mathrm{TVaR}_t(X)dt > \int_p^1 \mathrm{TVaR}_t(Y)dt$ for some p in $(0, p_0]$, we must have $H_1(X) > H_1(Y)$, because $\mathrm{TVaR}_t(X) \geq \mathrm{TVaR}_t(Y)$ for t in $[0, p_0]$. This is a contradiction. Hence, (27) holds for all p in $[0,1]$ and, consequently, $X \leq_{tvar[2]} Y$. ■

Finally, we return to the example in Section 1 based on two Pareto random variables X and Y with parameters $\alpha_1 = 2.4, \beta_1 = 3$ and $\alpha_2 = 2, \beta_2 = 2$, respectively. Recall that $X \nleq_n Y$ for $n \geq 1$. Since $H_1(X) = 5.8163 < 6 = H_1(Y)$ and the TVaR curves of X and Y satisfy the assumptions of Theorem 22 (see Figure 1), it follows $X \leq_{tvar[2]} Y$ (observe in Figure 2 that the curve

$T_X^{[2]}(p)$ lies everywhere below $T_Y^{[2]}(p)$). We conclude that an agent with a distortion $g \in \widehat{D}_3$ prefers X to Y.

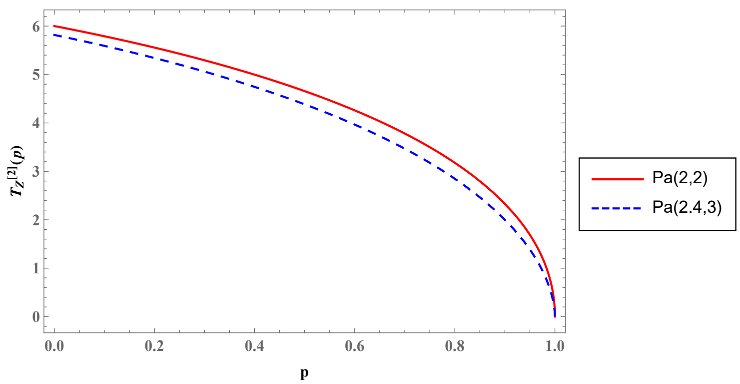

Figure 2: Curves $T_X^{[2]}(p)$ and $T_Y^{[2]}(p)$ as functions of p

7 Conclusions

In Yaari's model, people base their preferences on distortion risk measures of the form (1). In the context of the premium principle of Wang (1996), risk averse agents use concave distortions and the stop loss order characterizes their shared preferences. To increase the number of comparable risks, Wang and Young (1998) consider a sequence $\{D_n, n \geq 1\}$ of classes of progressively more risk averse agents using distortions whose successive derivatives alternate in sign and characterize their preferences by a sequence of partial orders introduced by Muliere and Scarsini (1989). Under Wang and Young's approach, if a risk X is preferred to a risk Y for a certain class D_n, then necessarily $E(X) \leq E(Y)$. In this paper, we have characterized, by a sequence of distortion-free partial orders based on comparing successive integrals of TVaR curves, the risk preferences of classes $\{\widehat{D}_n \subset D_n, n \geq 3\}$ of agents exclusively concerned with tail risks, irrespective of their mean values. Specifically, an agent with distortion $g \in \widehat{D}_n$ is more risk averse than an agent with the distortion $g_1(t)$ given by (3), where risk aversion is measured by the nth degree coefficient of dual risk aversion (7). As the main contributions, our approach (1) increases the completeness of risk orderings by comparing risks that are not comparable under Wang and Young's approach and (2) provides a plausible economic interpretation of the new orderings. Of course, our approach also has limitations. One is that the choice of g_1 as a benchmark to compare risk aversion may be considered somewhat arbitrary. Naturally, the choice is motivated by the role played by g_1 in interpreting the area under the TVaR curve. This suggests that another distortion g_2 may similarly lead to another sequence of partial orderings that can be a topic for future research.

8 Appendix A. Some proofs.

Proof of Theorem 3. Let h be a nth degree dual risk averse distortion function. First, note that

$$\int_0^\infty h(\bar{F}(x))dx = \int_0^1 F^{-1}(t)d\tilde{h}(t), \tag{28}$$

where $\tilde{h}(t) = 1 - h(1-t)$. The hypothesis of the dual theory in our context is that agents will choose among random variables to minimize (28). Given X and Y, we can write

$$\begin{aligned}
&\int_0^\infty h(\bar{G}(x))dx - \int_0^\infty h(\bar{F}(x))dx \\
=&\int_0^1 (G^{-1}(t) - F^{-1}(t))d\tilde{h}(t) \\
=&\sum_{k=1}^{n-1} \tilde{h}^{(k)}(0)(G_{k+1}^{-1}(0) - F_{k+1}^{-1}(0)) + \int_0^1 (G_n^{-1}(t) - F_n^{-1}(t))\tilde{h}^{(n)}(t)dt \\
=&\sum_{k=1}^{n-1}(-1)^{k+1}h^{(k)}(1)(G_{k+1}^{-1}(0) - F_{k+1}^{-1}(0)) \\
&+\int_0^1 (G_n^{-1}(t) - F_n^{-1}(t))(-1)^{n+1}h^{(n)}(1-t)dt.
\end{aligned}$$

A nth dual risk averter prefers X to Y if this difference is positive, which happens whenever conditions (a) and (b) in Definition 1 hold. Conversely, if every nth degree dual risk averter prefers X to Y, then necessarily Y has more nth degree dual risk than X (otherwise, as $h^{(k)}$ has no restrictions for $k = 1, ..., n-1$, we can give examples that contradict the claim). ■

Proof of Theorem 10. The proof follows the same steps as the proof of Theorem 3 of Jindapon and Neilson (2007). To prove the sufficient condition, rescale the distortion i so that $i^{(n-1)}(p_0) = j^{(n-1)}(p_0)$, where p_0 is the same as in Definition 9(b) and note that the solution of $U_i(t)$ minimization is the same as the solution from minimizing[17] $\frac{U_i(t)}{\tilde{i}^{(n-1)}(p_0)}$, where $\tilde{i}(p) = 1 - i(1-p)$ for all $p \in (0,1)$. From the comonotonicity of X and Y we can write

$$\begin{aligned}
U_i(t) &= \int_0^1 M_t^{-1}(p)d\tilde{i}(p) + c(t) \\
&= \int_0^1 (tF^{-1}(p) + (1-t)G^{-1}(p))d\tilde{i}(p) + c(t).
\end{aligned}$$

[17]By assumptions, we have $(-1)^n i^{(n-1)}(p_0) \geq 0$. Therefore, $\tilde{i}^{(n-1)}(p_0) \geq 0$.

The first order condition on $\frac{U_i(t)}{\tilde{i}^{(n-1)}(p_0)}$ yields

$$\frac{c'(t_i)}{\tilde{i}^{(n-1)}(p_0)} = \frac{\int_0^1 (G^{-1}(p) - F^{-1}(p))d\tilde{i}(p)}{\tilde{i}^{(n-1)}(p_0)}.$$

Define

$$\theta = \frac{\int_0^1 (G^{-1}(p) - F^{-1}(p))d\tilde{i}(p)}{\tilde{i}^{(n-1)}(p_0)} - \left(\frac{\int_0^1 (G^{-1}(p) - F^{-1}(p))d\tilde{j}(p)}{\tilde{j}^{(n-1)}(p_0)}\right).$$

Integration by parts yields

$$\theta = \int_0^1 \left[(F_{n-1}^{-1}(p) - G_{n-1}^{-1}(p))\right] \left[\frac{\tilde{j}^{(n-1)}(p)}{\tilde{j}^{(n-1)}(p_0)} - \frac{\tilde{i}^{(n-1)}(p)}{\tilde{i}^{(n-1)}(p_0)}\right] dp.$$

From j is more nth degree dual risk averse than i, it follows

$$\frac{\tilde{j}^{(n)}(p)}{\tilde{j}^{(n-1)}(p)} \geq \frac{\tilde{i}^{(n)}(p)}{\tilde{i}^{(n-1)}(p)}, \quad p \in (0,1). \tag{29}$$

Since $\tilde{j}^{(n-1)}(p)$ and $\tilde{i}^{(n-1)}(p)$ have the same sign, we can write

$$\frac{d}{dp}\log\left[\frac{\tilde{j}^{(n-1)}(p)}{\tilde{i}^{(n-1)}(p)}\right] = \frac{\tilde{j}^{(n)}(p)}{\tilde{j}^{(n-1)}(p)} - \frac{\tilde{i}^{(n)}(p)}{\tilde{i}^{(n-1)}(p)}. \tag{30}$$

Combining (29) and (30), we see that $\frac{\tilde{j}^{(n-1)}(p)}{\tilde{i}^{(n-1)}(p)}$ is non-decreasing. Then, $\frac{\tilde{j}^{(n-1)}(p)}{\tilde{j}^{(n-1)}(p_0)} \leq \frac{\tilde{i}^{(n-1)}(p)}{\tilde{i}^{(n-1)}(p_0)}$ for $p \leq p_0$ and $\frac{\tilde{j}^{(n-1)}(p)}{\tilde{j}^{(n-1)}(p_0)} \geq \frac{\tilde{i}^{(n-1)}(p)}{\tilde{i}^{(n-1)}(p_0)}$ for $p \geq p_0$. Since $G_{n-1}^{-1}(p)$ and $F_{n-1}^{-1}(p)$ cross only once at p_0, as indicated in Definition 9(b), then $\theta \leq 0$. Since $i^{(n-1)}(p_0) = j^{(n-1)}(p_0)$, $c'(t_i) \leq c'(t_j)$ and hence $t_i \leq t_j$, because c is convex.

To prove the converse, suppose that there exists some $p_0 \in (0,1)$ such that $\frac{\tilde{j}^{(n)}(p_0)}{\tilde{j}^{(n-1)}(p_0)} < \frac{\tilde{i}^{(n)}(p_0)}{\tilde{i}^{(n-1)}(p_0)}$. Because i and j are infinitely continuously differentiable, there exists a neighborhood E_0 of p_0 such that $\frac{\tilde{j}^{(n)}(p)}{\tilde{j}^{(n-1)}(p)} < \frac{\tilde{i}^{(n)}(p)}{\tilde{i}^{(n-1)}(p)}$ for all $p \in E_0$. It follows from (30) that $\frac{\tilde{j}^{(n-1)}(p)}{\tilde{i}^{(n-1)}(p)}$ is non-increasing. Then, $\frac{\tilde{j}^{(n-1)}(p)}{\tilde{j}^{(n-1)}(p_0)} \geq \frac{\tilde{i}^{(n-1)}(p)}{\tilde{i}^{(n-1)}(p_0)}$ for $p \leq p_0$ and $\frac{\tilde{j}^{(n-1)}(p)}{\tilde{j}^{(n-1)}(p_0)} \leq \frac{\tilde{i}^{(n-1)}(p)}{\tilde{i}^{(n-1)}(p_0)}$ for $p \geq p_0$. Now choose X and Y such that $G_{n-1}^{-1}(p) - F_{n-1}^{-1}(p)$ is a function with support in E_0 and Y differs from X by a simple increase in nth degree dual risk with a crossing at $p = p_0$. Then, for all $p \in E_0$ such that $p \leq p_0$, $G_{n-1}^{-1}(p) \leq F_{n-1}^{-1}(p)$ and $G_{n-1}^{-1}(p) \geq F_{n-1}^{-1}(p)$ for all $p \geq p_0$. Then $\theta \geq 0$. It follows that $t_i \geq t_j$, a contradiction. ■

Acknowledgement

We acknowledge support received from the Ministerio de Ciencia e Innovación (Spain) under grant PID2020-116216GB-I00.

References

[1] Arrow, K.J., 1970. Essays in the Theory of Risk Bearing. North-Holland, Amsterdam.

[2] Artzner, P., Delbaen, F., Eber, J. M., Heath, D., 1999. Coherent measures of risk. Math. Finance 9(3), 203-228.

[3] Caballé, J., Pomansky, A., 1996. Mixed risk aversion. J. Econ. Theory 71(2), 485-513.

[4] Castaño-Martínez, A., López-Blázquez, F., Pigueiras, G., Sordo, M.A., 2020. A method for constructing and interpreting some weighted premium principles. ASTIN Bull. 50, 1037-1064.

[5] Chateauneuf, A., Gajdos, T., Wilthien, P.H., 2002. The principle of strong diminishing transfer. J. Econ. Theory 103, 311–333.

[6] Chiu, W.H., 2005. Skewness preference, risk aversion, and the precedence relations on stochastic changes. Manage. Sci. 51(12), 1816-1828.

[7] Denneberg, D., 1990. Premium calculation: Why standard deviation should be replaced by absolute deviation. ASTIN Bull. 20, 181-190.

[8] Denuit, M., Dhaene, J., Goovaerts, M., Kaas, R., 2005. Actuarial Theory for Dependent Risks. John Wiley & Sons, New York.

[9] Denuit, M., Eeckhoudt, L., 2010. A general index of absolute risk attitude. Manage. Sci. 56(4), 712-715.

[10] Eeckhoudt, L., Laeven, R., Schlesinger, H., 2020. Risk apportionment: the dual story. J. Econ. Theory 185, 1–27.

[11] Eeckhoudt, L., Laeven, R., 2021. Dual moments and risk attitudes. Oper. Res. 70, 1330-1341.

[12] Eeckhoudt, L., Schlesinger, H., 2006. Putting risk in its proper place. Am. Econ. Rev. 96(1), 280–289.

[13] Ekern, S., 1980. Increasing nth degree risk. Econ. Lett. 6, 329-333.

[14] Hu, T., Chen, O., 2020. On a family of coherent measures of variability. Insur. Math. Econ. 95, 173-182.

[15] Jindapon, P., Neilson, W.S., 2007. Higher-order generalizations of Arrow–Pratt and Ross risk aversion: A comparative statics approach. J. Econ. Theory 136(1), 719-728.

[16] Kimball, M.S., 1990. Precautionary saving in the small and in the large. Econometrica 58, 53-73.

[17] Menezes, C., Geiss, C., Tressler, J., 1980. Increasing downside risk. Amer. Economic Rev. 70, 921-932.

[18] Muliere, P., Scarsini, M., 1989. A note on stochastic dominance and inequality measures. J. Econ. Theory 49, 314-323.

[19] Pflug, G.C., Römisch, W., 2007. Modeling, Measuring and Managing Risk. World Scientific Books.

[20] Pratt, J.W., 1964. Risk aversion in the small and in the large. Econometrica 32, 122-136.
[21] Psarrakos, G., Sordo, M.A., 2019. On a family of risk measures based on proportional hazards models and tail probabilities. Insur. Math. Econ. 86, 232-240.
[22] Rao, M., Chen, Y., Vemuri, B.C., Wang, F., 2004. Cumulative residual entropy: A new measure of information. IEEE Trans. Inform. Theory 50, 1220–1228.
[23] Rockafellar, R.T., Uryasev, S., Zabarankin, M., 2006. Generalized deviations in risk analysis. Finance Stoch. 10, 51–74.
[24] Rothschild, M., Stiglitz, J., 1970. Increasing Risk: I. A definition. J. Econ. Theory 2, 225-243.
[25] Shaked, M., Shanthikumar, J.G., 2007. Stochastic Orders. In: Springer Series in Statistics, Springer.
[26] Sordo, M.A., Castaño-Martínez, A., Pigueiras, G., 2016. A family of premium principles based on mixtures of TVaRs. Insur. Math. Econ. 70, 397-405.
[27] Sordo, M.A., Psarrakos, G., 2017. Stochastic comparisons of inter-failure times under a common replacement policy. J. Appl. Probab. 54, 134-145.
[28] Sun, H., Chen, Y., Hu, T., 2022. Statistical inference for tail-based cumulative residual entropy. Insur. Math. Econ. 103, 66-95.
[29] von Neumann, J., Morgenstern, O., 1944. Theory of Games and Economic Behavior. Princeton University Press, Princeton.
[30] Wang, S., 1995. Insurance pricing and increased limits ratemaking by proportional hazards transforms. Insur. Math. Econ. 17, 43-54.
[31] Wang, S., 1996. Premium calculation by transforming the layer premium density. ASTIN Bull. 26, 71–92.
[32] Wang, S., 1998. An actuarial index of the right-tail risk. N. Am. Actuar. J. 2, 88-101.
[33] Wang, S., Young, V.R., 1998. Ordering risks: Expected utility theory versus Yaari's dual theory of risk. Insur. Math. Econ. 22, 145-161.
[34] Wei, W., 2017. Joint stochastic orders of high degrees and their applications in portfolio selections. Insur. Math. Econ. 76, 141-148.
[35] Yaari, M., 1986. Univariate and multivariate comparisons of risk aversion: a new approach. Cambridge University Press, Cambridge.
[36] Yaari, M., 1987. The dual theory of choice under risk. Econometrica 55, 95-115.