




Article

Stochastic Comparisons of Some Distances between Random Variables

Patricia Ortega-Jiménez , Miguel A. Sordo * and Alfonso Suárez-Llorens 

Departamento de Estadística e I. O., Facultad de Ciencias, Universidad de Cádiz, 11002 Cádiz, Spain; patricia.ortega@uca.es (P.O.-J.); alfonso.suarez@uca.es (A.S.-L.)

* Correspondence: mangel.sordo@uca.es; Tel.: +34-956-012836

Abstract: The aim of this paper is twofold. First, we show that the expectation of the absolute value of the difference between two copies, not necessarily independent, of a random variable is a measure of its variability in the sense of Bickel and Lehmann (1979). Moreover, if the two copies are negatively dependent through stochastic ordering, this measure is subadditive. The second purpose of this paper is to provide sufficient conditions for comparing several distances between pairs of random variables (with possibly different distribution functions) in terms of various stochastic orderings. Applications in actuarial and financial risk management are given.

Keywords: stochastic order; copula; distance; variability measure; premium principle



Citation: Ortega-Jiménez, P.; Sordo, M.A.; Suárez-Llorens, A. Stochastic Comparisons of Some Distances between Random Variables. *Mathematics* **2021**, *9*, 981. <https://doi.org/10.3390/math9090981>

Academic Editors: Antonio Di Crescenzo and Elvira Di Nardo

Received: 26 March 2021

Accepted: 23 April 2021

Published: 27 April 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Given a bivariate random vector $X = (X_1, X_2)$ with joint distribution function $K_X(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$ and marginal distribution functions $F_1(x) = P[X_1 \leq x]$ and $F_2(x) = P[X_2 \leq x]$, the random variable $|X_1 - X_2|$ describes the distance between X_1 and X_2 in a sense that depends on the dependence structure of the vector. Different structures assign different meanings to this random variable and lead, obviously, to different ways of computing the expectation $E(|X_1 - X_2|)$. The distance can be applied to random variables with identical and non-identical distribution functions, and we consider both cases. If $F_1 = F_2 = F$, then X_1 and X_2 are copies of the same random variable X with distribution function $F(x)$, and $E(|X_1 - X_2|)$ reveals information about X . An example is the case of independent and identically distributed random variables, in which $E(|X_1 - X_2|)$ is the Gini's mean difference of X_1 , a well-known measure of variability (see, for example, [1]). We show in this work that, when X_1 and X_2 are dependent copies of the same random variable X , $E(|X_1 - X_2|)$ is still a measure of variability of X . A purpose of this paper is to study the properties of this functional in a general setting, where X_1 and X_2 are not necessarily independent.

If X_1 and X_2 are independent (or, more generally, if they are linked by a symmetric dependence structure), $|X_1 - X_2|$ treats symmetrically the events $X_1 < X_2$ and $X_2 < X_1$. However, sometimes, it is convenient to use a characteristic of proximity by treating them differently (in finance, for example, an investor evaluates differently gains and losses). The random excess of X_1 over X_2 , $(X_1 - X_2)^+$, where $x^+ = \max\{x, 0\}$ denotes the positive part of x , is useful if we are interested in measuring the extent to which one random variable exceeds the other, rather than the distance between them in a bidirectional sense. Note that the absolute value $|X_1 - X_2|$ can be split into two terms, each describing the excess of one random variable over the other, as follows:

$$|X_1 - X_2| = (X_1 - X_2)^+ + (X_2 - X_1)^+.$$

If X_1 and X_2 are copies of the same variable X , then $E((X_1 - X_2)^+)$ also reveals information about X . For example, if X_1 and X_2 are independent, $E((X_1 - X_2)^+)$ is Gini's mean semidifference.

In general, the functional $E(\phi(|X_1 - X_2|))$, where ϕ is a non-negative real function, has been largely studied in mathematics literature, mainly in the context of the Monge–Kantorovich problem (see [2], and references therein). The interest on this and other functionals used to measure the degree of difference between two random quantities goes back at least to the 1930s and the important contributions by Gini (see [3]) and Hoeffding [4]. Given two random vectors $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$, another purpose of this paper is to find conditions under which

$$E(\Phi(|X_1 - X_2|)) \leq E(\Phi(|Y_1 - Y_2|)), \text{ for all } \Phi \in \Omega, \quad (1)$$

where Ω is a subset of increasing real functions. Different choices of Ω give rise to different stochastic orderings between $|X_1 - X_2|$ and $|Y_1 - Y_2|$. This problem was addressed in [5–7] for the case where \mathbf{X} and \mathbf{Y} have independent components with the same marginal distribution functions (see Section 2 below for details). Here, we are concerned with two random vectors $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$, whose components are not necessarily independent nor are they required to have identical distribution functions. In this case, we explore conditions under which

$$|X_1 - X_2| \leq_{st, icx} |Y_1 - Y_2| \text{ and } (X_1 - X_2)^+ \leq_{st, icx} (Y_1 - Y_2)^+, \quad (2)$$

where \leq_{st} and \leq_{icx} are the usual stochastic order and the increasing convex order, respectively (these orders will be defined in Section 2 below).

This work is organized as follows. Section 2 contains preliminaries, such as definitions and background for the stochastic orders and dependence notions used in this paper, as well as a review of the properties that a variability measure should satisfy. In Section 3, given a random variable X with distribution function F , we show that any functional of the form $\nu(X) = E(|X_1 - X_2|)$, where X_1 and X_2 are two copies of X with any type of dependence structure, is a measure of variability of X . More generally, the distribution function F_1 of X_1 is allowed to be a distortion of F (we will explain the meaning of this below). In Section 4, given two random vectors $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$, we obtain conditions (both in terms of the marginals and the copulas) to make comparisons of the form (2). Section 5 contains two applications. In Section 5.1, we define a general class of premium principles based on the class of variability measures studied in Section 3. In Section 5.2, in the context of portfolio risk management, we assess the inclusion of a new asset in a portfolio by using the results obtained in Section 4. Finally, Section 6 contains conclusions.

Throughout this paper, given two random vectors $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$, we denote by F_1, F_2 and G_1, G_2 the respective marginal distribution functions. Given any other random variable Z , we denote by F_Z its distribution function.

2. Preliminaries

Let $\mathbf{X} = (X_1, X_2)$ be a random vector with joint distribution function K_X and marginal distribution functions F_1 and F_2 , respectively. According to the Sklar theorem, the joint distribution K_X can be written as

$$K_X(x, y) = C(F_1(x), F_2(y)),$$

where C is the copula of the random vector (X_1, X_2) , that is, the joint distribution function of the vector-copula $(F_1(X_1), F_2(X_2))$ (see [8]). If F_1 and F_2 are continuous, then C is unique. The copula contains the information about the structure of dependency of the random vector (X_1, X_2) . For every copula C and every (u, v) on $[0, 1]^2$, it is well-known that

$$W(u, v) \leq C(u, v) \leq M(u, v), \quad (3)$$

where the copulas $W(u, v) = \max(u + v - 1, 0)$ and $M(u, v) = \min(u, v)$ are the Fréchet–Hoeffding bounds. Random variables with copula M are called comonotonic and random variables with copula W are called countermonotonic.

The motivation for the study of the properties of $E(|X_1 - X_2|)$, where X_1 and X_2 are not necessarily independent, comes from the fact that some probability metrics and measures of variability that are sometimes better known under other expressions, take this form for different copulas between X_1 and X_2 . To give some examples, note that

$$E(|X_1 - X_2|) = \int_{-\infty}^{\infty} (F_1(x) + F_2(x) - 2C(F_1(x), F_2(x)))dx. \quad (4)$$

If X_1 and X_2 are two copies of a random variable X with distribution function $F(x)$, (4) becomes

$$E(|X_1 - X_2|) = 2 \int_{-\infty}^{\infty} (F(x) - C(F(x), F(x)))dx. \quad (5)$$

If X_1 and X_2 are two independent copies of X , then

$$E(|X_1 - X_2|) = 2(E(\max(X_1, X_2)) - E(X))$$

is the Gini's mean difference (GMD) of X , a well-known index of variability (see, for example, [1]). If X_1 and X_2 are comonotonic, then (see [9] or [10])

$$E(|X_1 - X_2|) = \int_0^1 |F_1^{-1}(u) - F_2^{-1}(u)|du = \int_{-\infty}^{\infty} |F_1(x) - F_2(x)|dx, \quad (6)$$

which is the Wasserstein distance, a well-known characteristic of proximity of two random variables (see [11]). If X_1 and X_2 are countermonotonic, then

$$E(|X_1 - X_2|) = \int_0^1 |F_1^{-1}(u) - F_2^{-1}(1 - u)|du \quad (7)$$

(see, for example, [2]). It is easy to see that, if X_1 and X_2 are two copies of X , (7) can be rewritten as

$$E(|X_1 - X_2|) = 2E(|X - m_X|),$$

where m_X is the median of X . This measure is twice the median absolute deviation (MAD), another popular measure of variability.

In view of the above examples, it is natural to ask whether $E(|X_1 - X_2|)$, where X_1 and X_2 are two copies of X with a copula C , fulfills the requirements to be considered as a measure of variability of X . Recall that a measure of variability ν is a map from the set of random variables to \mathbf{R} , such that given a random variable X , $\nu(X)$ quantifies the variability of X . Next, we list a set of properties that a measure of variability should reasonably satisfy (see, for example, [12] and references therein):

- (P0) Law invariance: if X and Y have the same distribution, then $\nu(X) = \nu(Y)$.
- (P1) Translation invariance: $\nu(X + k) = \nu(X)$ for all X and all constant k .
- (P2) Positive homogeneity: $\nu(0) = 0$ and $\nu(\lambda X) = \lambda \nu(X)$ for all X and all $\lambda > 0$.
- (P3) Non-negativity: $\nu(X) \geq 0$ for all X , with $\nu(X) = 0$ if X is degenerated at $c \in \mathbf{R}$.

Bickel and Lehmann [13] also require $\nu(X)$ to be consistent with the dispersive order. Recall that two random variables X and Y are ordered in the dispersive order if the difference between any two quantiles of X is smaller than the corresponding quantiles of Y , where the quantile function of a random variable X with distribution function F is defined by $F^{-1}(\alpha) = \inf\{x : F(x) \geq \alpha\}$, $\alpha \in (0, 1)$. The formal definition is as follows.

Definition 1. Given two random variables X and Y with distribution functions F and G , respectively, we say that X is smaller than Y in the dispersive order (denoted by $X \leq_{disp} Y$) if $F^{-1}(p) - F^{-1}(q) \leq G^{-1}(p) - G^{-1}(q)$ for all $0 \leq q < p \leq 1$.

A functional ν satisfying properties (P0) to (P3) is said to be a measure of variability or spread in the sense of Bickel and Lehmann if it satisfies in addition (see [13]):

(P4) Consistency with dispersive order: if $X \leq_{disp} Y$, then $\nu(X) \leq \nu(Y)$.

A measure of variability in the sense of Bickel and Lehmann considers the variability or spread of a random variable throughout its distribution. Sometimes, however, there is an interest in measuring only the variability of X along the right tail of its distribution (in risk theory, for example, some popular measures focus on the variability of a risk X beyond the value at risk). When this is the case, the requirement on ν to be consistent with the dispersive order is too strong. A natural weaker requirement is to be consistent with the excess wealth order (see [14]), which is defined as follows.

Definition 2. Given two random variables X and Y with distribution functions F and G , respectively, we say that X is smaller than Y in the excess wealth order (denoted by $X \leq_{ew} Y$) if $\int_{F^{-1}(p)}^{\infty} \bar{F}(x) dx \leq \int_{G^{-1}(p)}^{\infty} \bar{G}(x) dx$, $\forall p \in (0, 1)$, where $\bar{F} = 1 - F$ and $\bar{G} = 1 - G$ are the tail (or survival) functions of X and Y , respectively.

This allows us to consider the following property.

(P5) Consistency with excess wealth order: if $X \leq_{ew} Y$, then $\nu(X) \leq \nu(Y)$.

Measures of variability have received great attention in the actuarial and financial literature (see [12,15–18], among others). In actuarial science, for example, a variability measure sometimes is combined with a location measure to build a premium principle (see [19]). For particular applications in this context, we may wish ν to satisfy the following properties:

(P6) Comonotonic additivity: if X and Y are comonotonic, then $\nu(X + Y) = \nu(X) + \nu(Y)$.

(P7) Subadditivity: $\nu(X + Y) \leq \nu(X) + \nu(Y)$ for all X and Y .

Furman et al. ([12]) say that ν is a coherent measure of variability if it satisfies (P0)–(P3) and (P7).

Next, we recall some other notions used in this paper. The sequence of inequalities (3) induces the following definition (see [8]).

Definition 3. Given two copulas C and C' , we say that C is smaller than C' in the concordance order (and write $C \prec C'$) if $C(u, v) \leq C'(u, v)$ for all $u, v \in (0, 1)$.

Obviously, $W \prec C \prec M$ for every copula C . The name of this order is due to the fact that some measures of concordance, such as Kendall's tau and Spearman's rho, are increasing with respect to \prec .

In Sections 4 and 5, we will make use of the following stochastic orders. The reader may consult the books [20–22] for properties and applications.

Definition 4. Let X and Y be two random variables with distribution functions F and G and finite expectations μ_X and μ_Y , respectively. Then, X is said to be smaller than Y :

- (i) in the usual stochastic order (denoted by $X \leq_{st} Y$) if $\bar{F}(t) \leq \bar{G}(t)$, for all t ,
- (ii) in the increasing convex order (denoted by $X \leq_{icx} Y$) if $\int_t^{\infty} \bar{F}(x) dx \leq \int_t^{\infty} \bar{G}(x) dx$, $\forall t$,
- (iii) in the convex order (denoted by $X \leq_{cx} Y$) if $E(X) = E(Y)$ and $X \leq_{icx} Y$,
- (iv) in the increasing concave order (denoted by $X \leq_{icv} Y$) if $\int_{-\infty}^t F(x) dx \geq \int_{-\infty}^t G(x) dx$, $\forall t$.

It can be shown that $X \leq_{st} Y$ (respectively $\leq_{cx}, \leq_{icx}, \leq_{icv}$) if and only if $E(\phi(X)) \leq E(\phi(Y))$ for all increasing (respectively convex, increasing convex, increasing concave) functions ϕ for which the expectations exist. When X_2 and Y_2 are independent copies of X_1 and Y_1 , respectively, it is well-known (see [5] and [6]) that

- $X_1 \leq_{disp} Y_1$ implies $|X_1 - X_2| \leq_{st} |Y_1 - Y_2|$,

- $X_1 \leq_{ew, cx} Y_1$ implies $|X_1 - X_2| \leq_{icx} |Y_1 - Y_2|$.

The result for the convex order was extended to the so-called s-convex order in [7]. In Sections 3 and 4, we extend these results to the case where X_2 and Y_2 are not necessarily independent from (nor are they required to have identical distribution functions as) X_1 and Y_1 , respectively. For this, we need the following notions (see [23,24]).

Definition 5. Let $\mathbf{X} = (X_1, X_2)$ be a random vector.

- We say that X_1 is stochastically increasing in X_2 , denoted by $X_1 \uparrow_{SI} X_2$, if $P[X_1 > x_1 \mid X_2 = x_2]$ is a nondecreasing function of x_2 for all x_1 .
- We say that $\mathbf{X} = (X_1, X_2)$ is positively dependent through stochastic ordering (PDS) if $X_1 \uparrow_{SI} X_2$ and $X_2 \uparrow_{SI} X_1$.

Intuitively, if $\mathbf{X} = (X_1, X_2)$ is PDS, then its components are more likely simultaneously to have large values, compared with a vector of independent random variables with the same marginal distributions. For relationships between this and other dependence notions see, for example, Table 2 in [25]. The negative dependence analog of Definition 5 is as follows (see [24]).

Definition 6. Let $\mathbf{X} = (X_1, X_2)$ be a random vector.

- We say that X_1 is stochastically decreasing in X_2 , denoted by $X_1 \downarrow_{SD} X_2$, if $P[X_1 > x_1 \mid X_2 = x_2]$ is a nonincreasing function of x_2 for all x_1 .
- We say that $\mathbf{X} = (X_1, X_2)$ is negatively dependent through stochastic ordering (NDS) if $X_1 \downarrow_{SD} X_2$ and $X_2 \downarrow_{SD} X_1$.

Intuitively, if $\mathbf{X} = (X_1, X_2)$ is NDS, one component of the vector will tend to be large when the other component is small. It is easy to see that a random vector \mathbf{X} with continuous marginals is PDS (respectively, NDS) if and only if $C(u, v)$ is componentwise concave (respectively, convex). It is also well-known (see [26]) that a continuous random vector \mathbf{X} with copula C has the property PDS (resp. NDS) if and only if its copula C is PDS (resp. NDS).

3. A Family of Measures of Variability

Let X_1 and X_2 be two random variables with respective continuous distribution functions F_1 and F_2 and finite expectations. If X_1 and X_2 are two independent copies of X , it is well-known (see [13]) that $v(X) = E(|X_1 - X_2|)$ is a measure of variability in the sense of Bickel and Lehmann (that is, it satisfies properties (P0) to (P4)). Let h be a distortion function, that is, a non-decreasing function from $[0, 1]$ to $[0, 1]$ such that $h(0) = 0$ and $h(1) = 1$ (given two distribution functions F and G , if $G = h \circ F$ we say that G is a distortion of F via h). In this section, we show that any functional of the form $E(|X_1 - X_2|)$, where $F_2 = F$ and $F_1 = h \circ F$, is a measure of variability of X . In particular, if h is the identity function ($h(t) = t$, for all $t \in [0, 1]$) and X_1 and X_2 have a NDS copula, this measure satisfies all the properties (P1 to P7) listed above.

The following theorem extends a result of [5], stated as Theorem 3.B.42 in the book [20], in two directions: first, we consider two random vectors with the same copula instead of two random vectors with independent components; and, second, we allow the first marginal of each vector to be a distortion of the other (via the same h) instead of taking two copies of the same random variable.

Theorem 7. Let X and Y be two random variables with distribution functions F and G , respectively and let h be a distortion function. Let $\mathbf{X} = (X_1, X_2)$ be a random vector with respective marginal distribution functions $F_1 = h \circ F$ and $F_2 = F$. Similarly, let $\mathbf{Y} = (Y_1, Y_2)$ be a random vector with marginal distribution functions $G_1 = h \circ G$ and $G_2 = G$, respectively. Suppose that \mathbf{X} and \mathbf{Y} have the same copula C . If $X \leq_{disp} Y$, then $|X_2 - X_1| \leq_{st} |Y_2 - Y_1|$.

Proof. Since the dispersive order is preserved by distortion functions (Theorem 13 in [27]), we have $X_1 \leq_{disp} Y_1$ and $X_2 \leq_{disp} Y_2$. Since \mathbf{X} and \mathbf{Y} have the same copula, it follows from Definition 2.1 in [28] and Theorem 1 in [29] that there exists a function Φ that maps stochastically (X_1, X_2) into (Y_1, Y_2) , i.e., $\Phi(X_1, X_2) =_{st} (Y_1, Y_2)$, defined as

$$\Phi(x_1, x_2) = (\Phi_1(x_1), \Phi_2(x_2)) = (G_1^{-1}(F_1(x_1)), G_2^{-1}(F_2(x_2))),$$

where $\Phi_i(\cdot)$, $i = 1, 2$, is an increasing function that satisfies

$$|\Phi_i(x) - \Phi_i(x')| \geq |x - x'|, \forall x, x' \in \mathbb{R}. \quad (8)$$

It follows from the assumptions that $\Phi_2(x) = \Phi_1(x) = G^{-1}F(x)$ for all x . Therefore,

$$\begin{aligned} |Y_2 - Y_1| &=_{st} |\Phi_2(X_2) - \Phi_1(X_1)|, \\ &=_{st} |\Phi_1(X_2) - \Phi_1(X_1)|, \\ &\geq_{st} |X_2 - X_1|, \end{aligned} \quad (9)$$

where the first and second equality in (9) follow from the fact that $\Phi(X_1, X_2) =_{st} (Y_1, Y_2)$ and $\Phi_2(\cdot) = \Phi_1(\cdot)$, respectively. The inequality follows from (8) by using Theorem 1.A.1 in [20]. \square

By taking $h(x) = x$ in Theorem 7, we have the following corollary.

Corollary 8. Let X_2 and Y_2 be two copies of X_1 and Y_1 , respectively, such that $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ have the same copula. If $X_1 \leq_{disp} Y_1$, then $|X_1 - X_2| \leq_{st} |Y_1 - Y_2|$.

Remark 9. Given two random vectors $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ with the same copula, the condition $X_i \leq_{disp} Y_i$, $i = 1, 2$ is equivalent to say that the bivariate random vectors \mathbf{X} and \mathbf{Y} are ordered in a multivariate dispersion sense, see [30].

Now, we can prove the following result.

Theorem 10. Let X be a random variable with strictly increasing distribution function F and let h be a strictly increasing distortion function. Let X_1 and X_2 be two random variables with copula C and marginal distribution functions F_1 and F_2 , respectively. Let $v_C(X) = E(|X_1 - X_2|)$.

- (i) If $F_1 = h \circ F$ and $F_2 = F$, then $v_C(X)$ is a comonotonic additive measure of variability in the sense of Bickel and Lehmann, that is, it satisfies properties (P0)–(P4) and (P6).
- (ii) If $F_1 = F_2 = F$ and the copula C is NDS, then $v_C(X)$ satisfies all the properties (P0) to (P7).

Proof. We first prove (i). Let C be the copula of X_1 and X_2 . From (4), we have

$$\begin{aligned} v_C(X) &= \int_{-\infty}^{\infty} (h(F(x)) + F(x) - 2C(h(F(x)), F(x))) dx \\ &= \int_0^1 (h(u) + u - 2C(h(u), u)) dF^{-1}(u). \end{aligned} \quad (10)$$

Clearly, $v_C(X) = 0$ if X is degenerated at $c \in \mathbf{R}$. This, together with the fact that F^{-1} is non-decreasing, $F_{X+k}^{-1}(x) = F_X^{-1}(x) + k$, for all k and $F_{\lambda X}^{-1}(x) = \lambda F_X^{-1}(x)$, for all $\lambda > 0$ (see [31]), ensures that $v_C(X)$ satisfies properties (P0) to (P3). Since, given two random variables Z_1 and Z_2 , the condition $Z_1 \leq_{st} Z_2$ implies that $E(Z_1) \leq E(Z_2)$, property (P4) (consistency of $v_C(X)$ with respect to the dispersive order) is a direct consequence of Theorem 7. Property (P6) follows from the fact that, if Z_1 and Z_2 are comonotonic, then $F_{Z_1+Z_2}^{-1}(u) = F_{Z_1}^{-1}(u) + F_{Z_2}^{-1}(u)$, for all $u \in (0, 1)$ (see [32]). Under the assumptions in (ii), we have

$$v_C(X) = 2 \int_0^1 (u - C(u, u)) dF^{-1}(u). \quad (11)$$

Standard arguments show that

$$\lim_{u \rightarrow i} F^{-1}(u)(u - C(u, u)) = 0, \quad i = 0, 1.$$

Therefore, integrating (11) by parts, we get

$$\nu_C(X) = 2 \int_0^1 F^{-1}(u) d(C(u, u) - u).$$

Since C is componentwise convex, (P5) follows from Theorem 8. (ii) in [33] and (P7) follows from Theorem 2.1 in [12]. \square

Example 11. Two functionals satisfying the assumptions of part (i) are $\nu_{C_1}(X) = \text{GMD}(X)$ and $\nu_{C_2}(X) = \int_{-\infty}^{\infty} |F_h(x) - F(x)| dx$, which is the Wasserstein distance between F and its distortion $F_h = h \circ F$, a variability measure introduced by [34]. Note that $\nu_{C_2}(X) = E(|X_1 - X_2|)$, where $F_1 = h \circ F$, $F_2 = F$ and C_2 is the Fréchet–Hoeffding upper bound copula (see (6)).

Example 12. Using (7), it follows from Theorem 10 (ii) that $\nu_C(X) = E(|X - m_X|)$, where m_X is the median of X , satisfies all the properties (P0) to (P7) listed above. This measure can be written in the form $\frac{1}{2} E(|X_1 - X_2|)$ where $F_1 = F_2 = F$ and where C is the Fréchet–Hoeffding lower bound copula (see (7) and the paragraph below it), which is an example of NDS copula (see [35] for this and other examples of NDS copulas).

4. Other Stochastic Comparisons

To begin this section, we consider two random vectors $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ with the same marginals. Denote by $\mathcal{R}(F_1, F_2)$ the space of bidimensional random vectors with marginal distribution functions F_1 and F_2 .

Theorem 13. Let $\mathbf{X} = (X_1, X_2) \in \mathcal{R}(F_1, F_2)$ and $\mathbf{Y} = (Y_1, Y_2) \in \mathcal{R}(F_1, F_2)$ be two random vectors with copulas C and C' , respectively. If $C' \prec C$, then:

- (i) $|X_1 - X_2| \leq_{icx} |Y_1 - Y_2|$.
- (ii) $(X_1 - X_2)^+ \leq_{icx} (Y_1 - Y_2)^+$.

Proof. Under the assumptions, it follows from Theorem 4 of [36] that $X_1 - X_2 \leq_{icx} Y_1 - Y_2$. This means that $E(\Phi(X_1 - X_2)) \leq E(\Phi(Y_1 - Y_2))$ for all increasing convex Φ . Since $\Phi(t) = \Psi(|t|)$ is increasing and convex for any increasing convex function Ψ , it follows that $E(\Psi|X_1 - X_2|) \leq E(\Psi|Y_1 - Y_2|)$ for all increasing and convex Ψ , which proves (i). The proof of (ii) is similar using the function $\Phi(t) = \Psi((t)^+)$. \square

Remark 14. An alternative proof of Theorem 13 can be given by using Theorem 1 in [37], which provides conditions to ensure, under the above assumptions, that $E(k(X_1, Y_2)) \geq E(k(Y_1, Y_2))$ for certain classes of functions k . The proof is based on proving that the functions $k(x, y) = \phi|x - y|$ and $k(x, y) = \phi(x - y)^+$, with ϕ increasing and convex, satisfy those conditions.

A more general type of comparison can be made between two random vectors with possibly different (but stochastically ordered) marginals. The following two results provide conditions to compare two random excesses. The first result is given in terms of the usual stochastic order and the second result in terms of the increasing convex order.

Theorem 15. Let $\mathbf{X} = (X_1, X_2)$ be a random vector with respective marginal distribution functions F_1 and F_2 . Similarly, let $\mathbf{Y} = (Y_1, Y_2)$ be a random vector with marginal distribution functions G_1 and G_2 , respectively. If \mathbf{X} and \mathbf{Y} have the same copula C , $X_1 \leq_{st} Y_1$ and $X_2 \geq_{st} Y_2$, then $(X_1 - X_2)^+ \leq_{st} (Y_1 - Y_2)^+$.

Proof. For $x \geq 0$,

$$\begin{aligned}\bar{F}_{(X_1-X_2)^+}(x) &= 1 - \int_{-\infty}^{+\infty} P[X_1 \leq F_2^{-1}(p) + x | X_2 = F_2^{-1}(p)] dp \\ &= 1 - \int_0^1 \partial_2 C(F_1(F_2^{-1}(p) + x), p) dp, \quad x > 0.\end{aligned}$$

Therefore, given $x \geq 0$,

$$\begin{aligned}\bar{G}_{(Y_1-Y_2)^+}(x) - \bar{F}_{(X_1-X_2)^+}(x) \\ = \int_0^1 (\partial_2 C(F_1(F_2^{-1}(p) + x), p) - \partial_2 C(G_1(G_2^{-1}(p) + x), p)) dp \geq 0,\end{aligned}$$

Since $G_1(x) \leq F_1(x)$ for all x , $G_2^{-1}(p) \leq F_2^{-1}(p)$ for all $p \in (0, 1)$ and $\partial_2 C$ increases in the first argument (since $\partial_2 C(u, p)$ is the distribution function of the random variable $(U|V = p)$). Therefore, $\bar{F}_{(X_1-X_2)^+}(x) \leq \bar{G}_{(Y_1-Y_2)^+}(x)$ for all $x \geq 0$, which ends the proof. \square

Theorem 16. Let $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ be two random vectors with copulas C and C' , and marginal distribution functions F_1, F_2 and G_1, G_2 , respectively. If C is NDS, $C' \prec C$, $X_1 \leq_{icx} Y_1$ and $X_2 \geq_{icv} Y_2$, then $(X_1 - X_2)^+ \leq_{icx} (Y_1 - Y_2)^+$.

Proof. Let us consider a vector $\mathbf{Y}^* = (Y_1^*, Y_2^*)$ with copula C and such that $Y_i^* =_{st} Y_i$ for $i = 1, 2$. From the assumptions, it follows that $X_2 \geq_{icv} Y_2^*$, which is equivalent to say that $-X_2 \leq_{icx} -Y_2^*$ (Theorem 4.A.1 in [20]). Since \mathbf{X} and \mathbf{Y}^* have the same copula C , the random vectors $\hat{\mathbf{X}} = (X_1, -X_2)$ and $\hat{\mathbf{Y}} = (Y_1^*, -Y_2^*)$ have the same copula $\hat{C}(u, v) = u - C(u, 1 - v)$. Moreover, since C is NDS (that is, componentwise convex), then $\hat{C}(u, v)$ is PDS (that is, componentwise concave). It follows from Corollary 2.7 in [38] that $X_1 - X_2 \leq_{icx} Y_1^* - Y_2^*$. Since $C' \prec C$, it follows from Theorem 13 that $Y_1^* - Y_2^* \leq_{icx} Y_1 - Y_2$. The result follows by using the fact that the increasing convex order is transitive and is preserved by the increasing convex transformation $\phi(t) = t^+$ (see Theorem 4.A.8(a) in [20]). \square

Remark 17. In particular, Theorem 16 holds when $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ have the same NDS copula C . In this case, $X_1 \leq_{icx} Y_1$ and $X_2 \geq_{icv} Y_2$ imply $(X_1 - X_2)^+ \leq_{icx} (Y_1 - Y_2)^+$.

Lemma 18. Let X and Y be two random variables that are symmetric about 0. Then:

- (i) If $X^+ \leq_{st} Y^+$, then $|X| \leq_{st} |Y|$.
- (ii) If $X^+ \leq_{icx} Y^+$, then $|X| \leq_{icx} |Y|$.

Proof. Let \bar{F}_{X^+} and $\bar{F}_{|X|}$ be the tail functions of X^+ and $|X|$, respectively. If X and Y are symmetric about 0, it is easy to see that $\bar{F}_{|X|}(t) = h(\bar{F}_{X^+}(t))$ for all t , where h is the concave distortion function

$$h(t) = \begin{cases} 2t & \text{if } 0 \leq t \leq 1/2 \\ 1 & \text{if } 1/2 < t \leq 1. \end{cases}$$

Now (i) and (ii) follow, respectively, from Theorem 2.6 (i) and Theorem 2.6 (v) in [39]. \square

The following result follows immediately from Theorem 15 and Lemma 18.

Corollary 19. Let $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ be two random vectors with the same copula C and with marginal distribution functions F_1, F_2 and G_1, G_2 , respectively. If $X_1 \leq_{st} Y_1$ and $X_2 \geq_{st} Y_2$, then $|X_1 - X_2| \leq_{st} |Y_1 - Y_2|$.

The following result is also an immediate corollary of Theorem 16 and Lemma 18.

Corollary 20. Let $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ be two random vectors with symmetric copulas C and C' , and marginal distribution functions F_1, F_2 and G_1, G_2 , respectively. If C is NDS, $C' \prec C$, $X_1 \leq_{icx} Y_1$ and $X_2 \geq_{icv} Y_2$, then $|X_1 - X_2| \leq_{icx} |Y_1 - Y_2|$.

Remark 21. In particular, Corollary 20 holds when $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ have the same symmetric NDS copula C . When this is the case, $X_1 \leq_{icx} Y_1$ and $X_2 \geq_{icv} Y_2$ imply $|X_1 - X_2| \leq_{icx} |Y_1 - Y_2|$.

Since the independence copula is both NDS and PDS, a particular case of Corollaries 19 and 20 is the following.

Corollary 22. Let $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ be two random vectors with independent components and with marginal distribution functions F_1, F_2 and G_1, G_2 , respectively.

- (i) If $X_1 \leq_{st} Y_1$ and $X_2 \geq_{st} Y_2$, then $|X_1 - X_2| \leq_{st} |Y_1 - Y_2|$.
- (ii) If $X_1 \leq_{icx} Y_1$ and $X_2 \geq_{icv} Y_2$, then $|X_1 - X_2| \leq_{icx} |Y_1 - Y_2|$.

The following corollaries extend Lemma 2.2 in [6] from the case of two random vectors with independent components to the case of two random vectors with the same symmetric NDS copula.

Corollary 23. Let $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ be two random vectors with the same symmetric NDS copula C and with marginal distribution functions F_1, F_2 and G_1, G_2 , respectively. If $X_1 \leq_{cx} Y_1$ and $X_2 \leq_{cx} Y_2$, then $|X_1 - X_2| \leq_{icx} |Y_1 - Y_2|$.

Proof. The assumption $X_1 \leq_{cx} Y_1$ implies $X_1 \leq_{icx} Y_1$. Since $X_2 \leq_{cx} Y_2$ holds if and only if $-X_2 \leq_{cx} -Y_2$ (Theorem 3.A.12 in [20]), it follows $-X_2 \leq_{icx} -Y_2$. This is equivalent to write $X_2 \geq_{icv} Y_2$; therefore, the result follows from Corollary 20. \square

Corollary 24. Let $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$ be two random vectors with the same symmetric NDS copula, such that $X_2 =_{st} X_1$ and $Y_1 =_{st} Y_2$, all variables having finite means. If $X_1 \leq_{ew} Y_1$, then $|X_1 - X_2| \leq_{icx} |Y_1 - Y_2|$.

Proof. It is well-known (see (3.C.7) in [20]) that $X_1 \leq_{ew} Y_1$ implies $X_1 - E(X_1) \leq_{cx} Y_1 - E(Y_1)$. The result follows from Corollary 23. \square

5. Applications

5.1. An Application in Actuarial Science

In actuarial theory, a premium principle is a decision rule used by the insurer in order to determine the price for a risk to be insured. More formally, given a random variable X describing an insurance risk, a premium principle T assigns to X a number $T(X)$ which is the premium to be charged for accepting the risk X (see [19] for an overview). The simplest premium principle is the net premium $T(X) = E(X)$, which does not load for risk. More general premium principles are obtained by adding a load to the net premium that reflects the danger associated with the risk. Since the danger is often interpreted in terms of variability, a number of premium principles are obtained by adding to the net premium a risk load proportional to a specific measure of variability. Examples include the following:

- $T_1(X) = E(X) + \lambda \sqrt{\text{Var}[X]}$, with $\lambda > 0$ (standard deviation premium principle)
- $T_2(X) = E(X) + \lambda \text{GMD}(X)$, with $\lambda > 0$ (Gini's premium principle)
- $T_3(X) = E(X) + \lambda E(|X - m_X|)$, with $\lambda > 0$ (Denneberg's premium principle)
- $T_4(X) = E(X) + \int_{-\infty}^{\infty} |F_h(x) - F(x)|^p dx$, where $p \geq 0$ and h is a distortion function (see [34,40]).

Other examples can be found in [41]. Following this schema, we define a general class of premium principles based on distances between random variables.

Definition 25. Given a risk X with distribution function F , let \mathfrak{F} be the family of premium principles of the form

$$T_C(X) = E(X) + \lambda E(|X_1 - X_2|), \text{ for some } \lambda > 0,$$

where X_1 and X_2 are two copies of X such that X_1 and X_2 have copula C .

From the results in Section 2, we see that T_2, T_3 and T_4 (for $p = 1$) are premium principles that belong to the family \mathfrak{F} for different choices of the copula C . Moreover, it follows from Theorem 10 that a premium principle $T_C \in \mathfrak{F}$ satisfies the following properties:

- (a) Risk loading: $T_C(X) \geq E(X)$.
- (b) Non unjustified risk loading: If $X = k \geq 0$ (k constant), then $T_C(X) = k$.
- (c) Translation invariance: $T_C(X + k) = T_C(X) + k$ for all constant k .
- (d) Scale invariance: $T_C(bX) = bT_C(X)$ for all constant $b > 0$.
- (e) Comonotonic additivity: if X and Y are comonotonic, then $T_C(X + Y) = T_C(X) + T_C(Y)$.
- (f) Subadditivity: if C is NDS, then $T_C(X + Y) \leq T_C(X) + T_C(Y)$ for all X and Y .

A premium principle satisfying the above properties that does not follow the schema $T(X) = E(X) + \lambda D(X)$, where D is a measure of variability of X , is the distortion premium principle [32], defined by

$$T_h(X) = - \int_{-\infty}^0 (1 - h[\bar{F}(x)]) dx + \int_0^{\infty} h[\bar{F}(x)] dx,$$

where h is a concave distortion function. Our next result is related to a property of T_h . Recall that, given a random variable Z with distribution function F_Z , the family

$$\Pi_Z \equiv \{X = \mu + \sigma Z : \mu \in \mathbb{R}, \sigma > 0\}$$

is called a location-scale family of random variables. It is shown in [42] that the distortion premium principle T_h reduces to T_1 (the standard deviation premium principle) for location-scale families of distributions. Next, we give a similar result involving the premium principle T_C given in Definition 25.

Theorem 26. Consider a location scale family Π_Z and let $T_C \in \mathfrak{F}$. Then, T_C reduces to T_1 (standard deviation premium principle) or, equivalently, to T_h (distortion premium principle) on Π_Z .

Proof. Since T_1 is a special case of T_h on location-scale families of distributions [42], it suffices to prove that T_C reduces to T_h on Π_Z . Let $X \in \Pi_Z$. Since $F_X(x) = F_Z(\frac{x-\mu}{\sigma})$ for all x , if X_1 and X_2 are two copies of X with copula C , we have

$$\begin{aligned} T_C(X) &= E(X) + \lambda E[|X_1 - X_2|] \\ &= \mu + \sigma E(Z) + 2\lambda \int_{-\infty}^{\infty} (F_X(x) - C(F_X(x), F_X(x))) dx \\ &= \mu + \sigma E(Z) + 2\lambda \int_{-\infty}^{\infty} (F_Z(z) - C(F_Z(z), F_Z(z))) \sigma dz \\ &= \mu + \sigma E(Z) + 2\lambda \sigma E[|Z_1 - Z_2|], \end{aligned}$$

where Z_1 and Z_2 are two copies of Z with copula C (here, we have used that copulas are invariant for strictly monotone transformations of the random variables). Now, we equate this expression with $T_h(X)$ to obtain

$$\lambda = \frac{T_h(Z) - E(Z)}{E[|Z_1 - Z_2|]}$$

where we have used that $T_h(X) = \mu + \sigma T_h(Z)$, because T_h is scale and translation invariant. Observe that λ is independent of μ and σ ; therefore, we conclude that T_C reduces to T_h on C . \square

5.2. An Application in Portfolio Risk Management

In portfolio risk management, investors diversify portfolios to reduce market risk. On average, a portfolio with several assets exhibits, unless they are perfectly correlated, less variability in returns than a portfolio with only one asset. To illustrate the results in Section 4, let us consider an investor whose portfolio has only one asset A with log-return X_2 . During the diversification process, the investor is concerned with the risk of two assets B and C with log-returns X_1 and Y_1 , respectively, that might be included in her/his portfolio. Here, the risk of these two assets B and C must be considered in relation to the asset A, which acts as a hedge. One way to assess the impact of each of these two assets is by comparing the distances $|X_1 - X_2|$ and $|Y_1 - X_2|$ in some stochastic sense. A smaller distance suggests a higher degree of similarity between the log-returns of the assets when evaluated jointly. If the assets B and C move in the opposite direction as the hedge A, a higher degree of similarity with A intuitively reduces the risk of the diversified portfolio. An alternative method to assess the impact of B and C on the portfolio is by using measures of contagious (see Section 4 in [43]).

Recall that the log return of an asset at week t is defined by $r_t = \log(p_t/p_{t-1})$, where p_t is the price of the asset at week t . For our empirical example, we work with log-returns of three stocks included in Nasdaq Composite index: Zoom Video Communications (X_1), Moderna (Y_1) and Booking Holdings Inc ($X_2 = Y_2$) (we have selected three companies that were affected differently at the beginning of the financial crisis of COVID-19). The study is based on samples of size $n = 64$ for each financial institution ($\{x_{1i}\}$, $\{y_{1i}\}$ and $\{x_{2i} = y_{2i}\}$, for $i = 1, \dots, 64$), measuring the share value from 23 December 2019 until 15 March 2021. Data were gathered from the public website <http://es.finance.yahoo.com>, accessed on 22 March 2021, and are related to the weekly close of trading to eliminate the time dependent effect. Suppose that, initially, the whole portfolio of our investor consists of stocks in only one company: Booking Holdings, Inc. To reduce risk, the investor plans to invest either in Zoom or Moderna and faces the problem of which of them should be chosen. A method that helps to make a decision, as explained above, is to compare the distances between the components of the random vectors $\mathbf{X} = (X_1, X_2)$ and $\mathbf{Y} = (Y_1, Y_2)$.

Figure 1 plots the empirical distribution function of sample absolute differences $|x_{1i} - x_{2i}|$ between Zoom and Booking (green curve) and $|y_{1i} - y_{2i}|$ between Moderna and Booking (blue curve), for $i = 1, \dots, 64$. The blue curve starts above; at some point around $x = 0.005$, it crosses from above to below the green curve and, after that point, it seems to be everywhere below. The graphic is consistent with a model where $|X_1 - X_2|$ and $|Y_1 - Y_2|$ are ordered in the increasing convex order. Next, we give statistical significance to this conclusion.

We first perform some tests to study the marginal distributions of X_1 , X_2 and Y_1 . In order to check randomness, the classical runs test is performed with p -values 0.6143, 0.3134 and 0.2077, respectively. Symmetry is tested using the symmetry test by [44], obtaining p -values 0.388, 0.81, and 0.174, respectively. The Kolmogorov–Smirnov test for normality gives, respectively, the p -values 0.5424, 0.9154, and 0.3486. Therefore, there is not significant evidence to reject the hypothesis that the three log return distributions are random, symmetric, and normal.

A unilateral F-test for paired data, performed for testing the hypothesis of equality of variances against $\sigma_{X_1} < \sigma_{Y_1}$, gives a p -value of 0.000493, showing significant evidence that $\sigma_{X_1} < \sigma_{Y_1}$. The p -value of the t -test for testing $\mu_{X_1} = \mu_{Y_1}$ against $\mu_{X_1} \neq \mu_{Y_1}$ is 0.7124, so we can not reject the equality of means. From the assumptions of normality, $\mu_{X_1} = \mu_{Y_1}$ and $\sigma_{X_1} < \sigma_{Y_1}$, it follows $X_1 \leq_{icx} Y_1$ (see Table 2.2 in [22]).

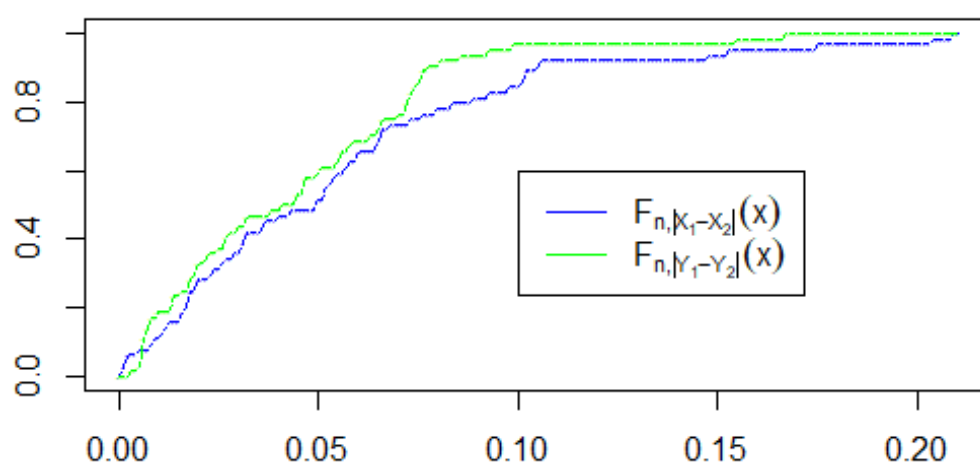


Figure 1. Distribution functions of empirical absolute differences. The green curve corresponds to $F_{n,|X_1-X_2|}(x)$ and the blue curve to $F_{n,|Y_1-Y_2|}(x)$.

The copulas C and C' are adjusted by using the goodness of fit test based on Kendall's process [45,46]. Considering a bivariate normal (BN) copula, we obtain p -values 0.74 and 0.79, respectively; therefore, there is not statistical evidence to reject that C and C' are BN. Since the bivariate normal copula parameter is the Pearson correlation coefficient ρ_X , we perform the Williams's Test (bilateral) [47,48] for testing the hypothesis $\rho_X = \rho_Y$ against $\rho_X \neq \rho_Y$ when the vectors share one component ($X_2 = Y_2$). The p -value, 0.4622, indicates that we cannot reject the equality, which leads us to admit that $C = C' \sim BN(\rho)$. Since the sample estimate of the Pearson's correlation coefficient is negative, we test $\rho = 0$ against $\rho < 0$ by running the test for association between paired samples using Pearson's correlation coefficient for the vector X . The p -value, 0.0442, suggests that $\rho < 0$, which means that the copulas are NDS (see Example 4.1 in [24]).

To conclude: the assumptions (1) $X_1 \leq_{icx} Y_1$ and that (2) the copulas C and C' are equal and are NDS, are supported by statistical significance. It follows from Corollary 20 that $|X_1 - X_2| \leq_{icx} |Y_1 - Y_2|$, which indicates that Zoom leads to a less risky portfolio than Moderna.

6. Conclusions

Given two random variables X_1 and X_2 that are not necessarily independent, we have provided several results concerning the distances $|X_1 - X_2|$, $(X_1 - X_2)^+$ and their expectations. The most remarkable results of this study can be summarized as follows:

- (a) If X is a random variable with strictly increasing distribution function F and X_1 and X_2 are two random variables with a NDS (negative dependent through stochastic ordering) copula C and with marginal distribution functions $F_1 = F_2 = F$, then $\nu(X) = E(|X_1 - X_2|)$ is a variability measure satisfying the following properties: law invariance, translation invariance, positive homogeneity, non-negativity, consistency with the dispersive order, consistency with the excess wealth order, comonotonic additivity, and subadditivity. An example is the median absolute deviation $\nu(X) = E(|X - m_X|)$, where m_X is the median of X , which can be written in the form $\frac{1}{2}E(|X_1 - X_2|)$, where C is the Fréchet–Hoeffding lower bound copula.
- (b) Given two random vectors (X_1, X_2) and (Y_1, Y_2) with possibly different marginals and copulas, we have given conditions, in terms of several stochastic orders, under which $|X_1 - X_2| \leq_{st,icx} |Y_1 - Y_2|$ and $(X_1 - X_2)^+ \leq_{st,icx} (Y_1 - Y_2)^+$.

Two applications have been provided. In actuarial science, given a risk X , we have proposed a general class of premium principles of the form $T_C(X) = E(X) + \lambda E(|X_1 - X_2|)$, for some $\lambda > 0$, where X_1 and X_2 are two copies of X with copula C . In portfolio risk management, we have assessed the inclusion of a new asset in a portfolio by comparing absolute values of differences. It is a question of future research to determine the circum-

stances under which the criterion used in Section 5.2 to include a new asset in a portfolio gives rise to a portfolio with smaller realized variance.

Finally, it is interesting to note that the random excess $(X_1 - X_2)^+$ also has an appealing role in the context of risk management and quantitative finance. If $X_1(t)$ and $X_2(t)$ are the prices of two risky assets at time t , the payoff of the option that gives the buyer the right to exchange the second asset for the first at the expiry time t (called exchange option) is $(X_1(t) - X_2(t))^+$. In this context, the results in this paper can be used to compare different payoffs in a similar manner as in Section 5.2.

Author Contributions: Investigation, P.O.-J., M.A.S., A.S.-L. All authors contributed equally to this work. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the 2014–2020 ERDF Operational Program and the Department of Economy, Knowledge, Business, and University of the Regional Government of Andalusia under grant FEDER-UCA18-107519. P.O.-J. acknowledges University of Cádiz for the PhD grant (call 2018) linked to the project MTM2017-89577-P financed by Ministerio de Economía y Competitividad of Spain.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yitzhaki, S. Gini's Mean difference: A superior measure of variability for non-normal distributions. *Metron* **2003**, *61*, 285–316.
2. Rachev, S.T. The Monge–Kantorovich Mass Transference Problem and Its Stochastic Applications. *Theory Probab. Appl.* **1985**, *29*, 647–676. [\[CrossRef\]](#)
3. Giorgi, G.M. Bibliographic portrait of the Gini concentration ratio. *Metron* **1990**, *48*, 183–221.
4. Hoeffding, W. Masstabvariante Korrelationstheorie. *Sehr. Math. Inst. Univ. Berl.* **1950**, *5*, 181–233.
5. Giovagnoli, A.; Wynn, H. Multivariate dispersion orderings. *Stat. Probab. Lett.* **1995**, *22*, 325–332. [\[CrossRef\]](#)
6. Kochar, S.C.; Carriere, K.C. Connections among various variability orderings. *Stat. Probab. Lett.* **1997**, *35*, 321–333. [\[CrossRef\]](#)
7. Bassan, B.; Denuit, M.; Scarsini, M. Variability orders and mean differences. *Stat. Probab. Lett.* **1999**, *45*, 121–130. [\[CrossRef\]](#)
8. Nelsen, R.B. *An Introduction to Copulas*; Lecture Notes in Statistics No. 139; Springer: New York, NY, USA, 1999.
9. Dall'Aglio, G. Sugli estremi dei momenti delle funzioni di ripartizione doppia. *Ann. Scuola Norm.-Sci.* **1956**, *10*, 35–74.
10. Vallender, S.S. Calculation of the Wasserstein Distance Between Probability Distributions on the Line. *Theory Probab. Appl.* **1972**, *18*, 784–786. [\[CrossRef\]](#)
11. Kantorovich, L.V.; Rubinstein, G.S. On a space of completely additive functions. *Vestn. Leningr. Univ.* **1958**, *13*, 52–59.
12. Furman, E.; Wang, R.; Zitikis, R. Gini-type measures of risk and variability: Gini shortfall, capital allocations, and heavy-tailed risks. *J. Bank Financ.* **2017**, *83*, 70–84. [\[CrossRef\]](#)
13. Bickel, P.J.; Lehmann, E.L. *Descriptive Statistics for Nonparametric Models (IV. Spread)*; Jureckova, J., Ed.; Academia: Prague, Czech Republic, 1979; pp. 33–40.
14. Sordo, M.A. Comparing tail variabilities of risks by means of the excess wealth order. *Insur. Math. Econ.* **2009**, *45*, 466–469. [\[CrossRef\]](#)
15. Sordo, M.A.; Suárez-Llorens, A. Stochastic comparisons of distorted variability measures. *Insur. Math. Econ.* **2011**, *49*, 11–17. [\[CrossRef\]](#)
16. Hu, T.; Chen, O. On a family of coherent measures of variability. *Insur. Math. Econ.* **2020**, *95*, 173–182. [\[CrossRef\]](#)
17. Rockafellar, R.T.; Uryasev, S.; Zabarankin, M. Generalized deviations in risk analysis. *Financ. Stoch.* **2006**, *10*, 51–74. [\[CrossRef\]](#)
18. Psarrakos, G.; Sordo, M.A. On a family of risk measures based on proportional hazards models and tail probabilities. *Insur. Math. Econ.* **2019**, *86*, 232–240. [\[CrossRef\]](#)
19. Young, V.R. Premium principles. In *Encyclopedia of Actuarial Science*; Wiley: New York, NY, USA, 2004; pp. 1322–1331.
20. Shaked, M.; Shanthikumar, J.G. *Stochastic Orders*; Springer: New York, NY, USA, 2007.
21. Müller, D.; Stoyan, D. *Comparison Methods for Stochastic Models and Risks*; Wiley: New York, NY, USA, 2002.
22. Belzunce, F.; Riquelme, C.M.; Mulero, J. *An Introduction to Stochastic Orders*; Academic Press: New York, NY, USA, 2015.
23. Barlow, R.E.; Proschan, F. *Statistical Theory of Reliability and Life Testing: Probability Models*; Rinehart and Winston; Holt: New York, NY, USA, 1981.
24. Block, H.W.; Savits, T.H.; Shaked, M. A concept of negative dependence using stochastic ordering. *Stat. Probab. Lett.* **1985**, *3*, 81–86. [\[CrossRef\]](#)
25. Navarro, J.; Pellerey, F.; Sordo, M.A. Weak dependence notions and their mutual relationships. *Mathematics* **2021**, *9*, 81. [\[CrossRef\]](#)

26. Cai, J.; Wei, W. On the invariant properties of notions of positive dependence and copulas under increasing transformations. *Insur. Math. Econ.* **2012**, *50*, 43–49. [\[CrossRef\]](#)
27. Sordo, M.A.; Suárez-Llorens, A.; Bello, A. Comparison of conditional distributions in portfolios of dependent risks. *Insur. Math. Econ.* **2015**, *61*, 62–69. [\[CrossRef\]](#)
28. Fernández-Ponce, J.M.; Suárez-Llorens, A. A multivariate dispersion ordering based on quantiles more widely separated. *J. Multivar. Anal.* **2003**, *85*, 40–53. [\[CrossRef\]](#)
29. Arias-Nicolás, J.P.; Fernández-Ponce, J.M.; Luque-Calvo, P.; Suárez-Llorens, A. Multivariate dispersion order and the notion of copula applied to the multivariate t-distribution. *Probab. Eng. Inform. Sci.* **2005**, *19*, 363–375. [\[CrossRef\]](#)
30. Belzunce, F.; Ruiz, J.M.; Suárez-Llorens, A. On multivariate dispersion orderings based on the standard construction. *Stat. Probab. Lett.* **2008**, *78*, 271–281. [\[CrossRef\]](#)
31. Parzen, E. Nonparametric Statistical Data Modeling. *J. Am. Stat. Assoc.* **1979**, *74*, 105–121. [\[CrossRef\]](#)
32. Wang, S. Premium Calculation by Transforming the Layer Premium Density. *ASTIN Bull.* **1996**, *26*, 71–92. [\[CrossRef\]](#)
33. Sordo, M.A. Characterizations of classes of risk measures by dispersive orders. *Insur. Math. Econ.* **2008**, *42*, 1028–1034. [\[CrossRef\]](#)
34. López-Díaz, M.; Sordo, M.A.; Suárez-Llorens, A. On the L_p -metric between a probability distribution and its distortion. *Insur. Math. Econ.* **2012**, *51*, 257–264. [\[CrossRef\]](#)
35. Li, C.; Li, X. Preservation of increasing convex/concave order under the formation of parallel/series system of dependent components. *Metrika* **2018**, *81*, 445–464. [\[CrossRef\]](#)
36. Müller, A. On the waiting times in queues with dependency between interarrival and service times. *Oper. Res. Lett.* **2000**, *26*, 43–47. [\[CrossRef\]](#)
37. Cambanis, S.; Simons, G.; Stout, W. Inequalities for $E_k(x, y)$ when the marginals are fixed. *Zeitschrift für Wahrscheinlichkeitstheorie Und Verwandte Gebiete* **1976**, *36*, 285–294. [\[CrossRef\]](#)
38. Balakrishnan, N.; Belzunce, F.; Sordo, M.A.; Suárez-Llorens, A. Increasing directionally convex orderings of random vectors having the same copula, and their use in comparing ordered data. *J. Multivar. Anal.* **2012**, *105*, 45–54. [\[CrossRef\]](#)
39. Navarro, J.; del Aguila, Y.; Sordo, M.A.; Suárez-Llorens, A. Stochastic ordering properties for systems with dependent identically distributed components. *Appl. Stoch. Models Bus. Ind.* **2013**, *29*, 264–278. [\[CrossRef\]](#)
40. Yang, J.; Zhuang, W.; Hu, T. L_p -metric under the location-independent risk ordering of random variables. *Insur. Math. Econ.* **2014**, *59*, 321–324. [\[CrossRef\]](#)
41. Sordo, M.A.; Castaño-Martínez, A.; Pigueiras, G. A family of premium principles based on mixtures of TVaRs. *Insur. Math. Econ.* **2016**, *70*, 397–405. [\[CrossRef\]](#)
42. Young, V. Discussion of Christofides' Conjecture Regarding Wang's Premium Principle. *ASTIN Bull.* **1999**, *2*, 191–195. [\[CrossRef\]](#)
43. Ortega-Jiménez, P.; Sordo, M.A.; Suárez-Llorens, A. Stochastic orders and multivariate measures of risk contagion. *Insur. Math. Econ.* **2021**, *96*, 199–207. [\[CrossRef\]](#)
44. Miao, W.; Gel, Y.R.; Gastwirth, J.L. A new test of symmetry about an unknown median. In *Random Walk, Sequential Analysis and Related Topics: A Festschrift in Honor of Yuan-Shih Chow*; World Scientific Publishing Co.: Singapore, 2006; pp. 199–214.
45. Wang, W.; Wells, M.T. Model selection and semiparametric inference for bivariate failure-time data. *J. Am. Stat. Assoc.* **2000**, *95*, 62–72. [\[CrossRef\]](#)
46. Genest, C.; Quessy, J.F.; Remillard, B. Goodness-of-fit Procedures for Copula Models Based on the Probability Integral Transformation. *Scand. J. Stat.* **2006**, *33*, 337–366. [\[CrossRef\]](#)
47. Steiger, J.H. Tests for comparing elements of a correlation matrix. *Psychol. Bull.* **1980**, *87*, 245–251. [\[CrossRef\]](#)
48. Williams, E.J. *Regression Analysis*; Wiley: New York, NY, USA, 1959.